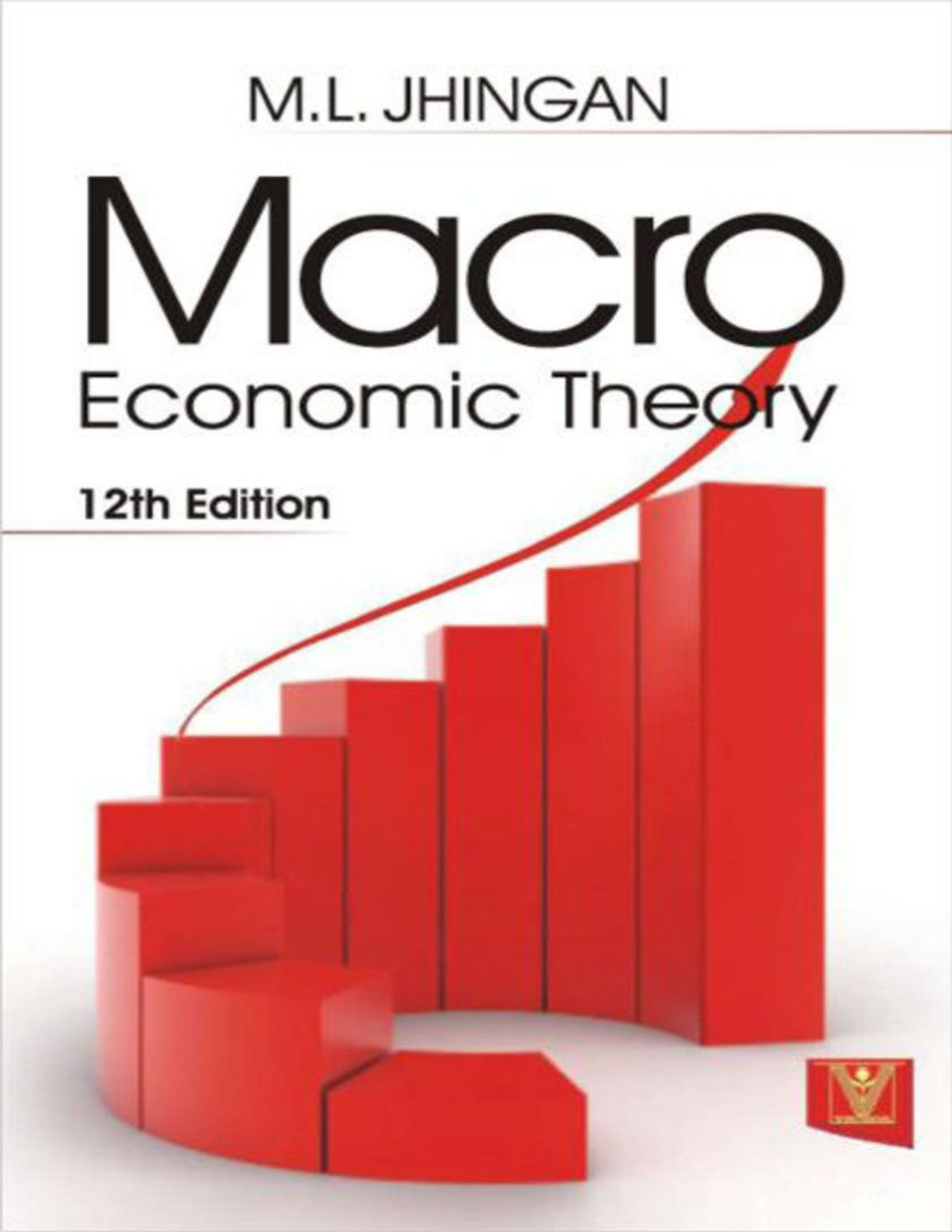
M.L. JHINGAN

Macro Economic Theory

12th Edition



MACROECONOMIC THEORY

By the same author :

THE ECONOMICS OF DEVELOPMENT & PLANNING

ADVANCED ECONOMIC THEORY

MODERN MICROECONOMICS

MONETARY ECONOMICS

INTERNATIONAL ECONOMICS

MONEY, BANKING AND INTERNATIONAL TRADE & PUBLIC FINANCE

MICROECONOMIC THEORY

PRINCIPLES OF KECONOMICS

PUBLIC FINANCE & INTERNATIONAL TRADE

DEMOGRAPHY

(Hindi Editions of all these Books are available)

(Punjabi Editions of the following Books are also available)

Modern Microeconomics
Macroeconomic Theory

MACROECONOMIC THEORY

12TH REVISED AND ENLARGED EDITION

M.L. JHINGAN

Retired Deputy Director,
Higher Education, Haryana

SIMPLY THE TEXT !

VRINDA PUBLICATIONS (P) LTD.

VRINDA PUBLICATIONS (P) LTD.
B-5, Ashish Complex (opp. Ahlcon Public School),
Mayur Vihar, Phase-I, Delhi-110091
Tel. : 22755315, 22755316 Fax : 22757220
E-mail : vrinda@ndf.vsnl.net.in
http:// www.vrindaindia.com

First Edition : 1983
Tenth Revised Edition : 1997
Reprinted Six times
Eleventh Revised & Enlaged Edition : 2003
Reprint : 2004, 2005, 2006, 2007, 2008
Twelfth Revised & Enlarged Edition : 2010
Reprint: 2011,2012

ISBN 978-81-8281-298-7
(Macroeconomic Theory – 12th Edition)

Typeset by : Harsimran Book Work & Graphics, Delhi
Printed at : Nisha Enterprises, Delhi

PREFACE

PREFACE TO THE TWELFTH EDITION

This is a substantially revised, enlarged and rearranged edition with new diagrams and in a new format.

Main Features

Nine Parts consisting of Introduction to Macroeconomics; National Income and Accounting; Macroeconomic Theory; Inflation and Business Cycles; Growth Models; Macroeconomic Policies; Modern Macroeconomics; Macroeconomics in Open Economy.

New Chapters : Complex Multipliers; The Saving Function (*Part Three*); The Monetarist Revolution (*Part Four*); The Solow-Swan Model; The Endogenous Growth Theory (*Part Six*); Extensions of IS-LM Models (*Part Seven*); The Real Business Cycle Theory; The New Keynesian Macroeconomics (*Part Eight*).

Rewrite of Chapters and Additions of Topics :

- The Classical Theory of Employment (Ch. 6).
- Theories of the Consumption Function : Consumption Puzzle; Drift Theory (Ch. 10)
- The Investment Function : The Present Value Criterion of Investment (Ch.11)
- Some New Theories of Investment : The Financial Theory of Investment; Tobin's *q* Theory of Investment (Ch. 16)
- The Demand for Money (Ch. 32)
- Theories of Interest Rate : Wicksell's and Fisher's Theories of Natural and Market Rate of Interest (Ch. 33)
- Inflation and Deflation : Bent Hansen's Excess Demand Model; Inflation as a Tax; Cost of Inflation (Ch. 37)
- Business Cycles : The Psychological Theory; The Cobweb Theory (Ch. 38)
- Macroeconomic Policy : Okun's Law; The Swan Model; The Assignment Problem; Rules vs. Discretion in Economic Policy; Expectations, Lucas Critique and New Classical Stabilisation Policy; Lags in Effects of Economic Policy; (Ch. 43)
- Macroeconomic Policy in Open Economy (Ch. 59 to 63)

I am grateful to many students and teachers from India and abroad who have favoured me with their valuable suggestions. My special thanks are to Mr. Chandar Kant, Post-graduate Department of Economics, Government College, Hisar (Haryana) who provided me useful material for revising the book.

I am sanguine that the present edition will prove more useful to candidates preparing for Hons., M.A.,

N.E.T., I.E.S., Civil Services and other competitive examinations.

Suggestions for further improvement of the book are solicited.

e-mail : vrinda@ndf.vsnl.net.in **M.L. JHINGAN**

BRIEF CONTENTS

PART FIVE – INFLATION AND BUSINESS CYCLES

37 Inflation and Deflation
38 Business Cycles

PART SIX – GROWTH MODELS

39 The Harrod-Domar Models
40 The Solow-Swan Model of Growth
41 The Solow-Swan Model of Economic Growth
42 The Endogenous Growth Theory
43 Steady State Growth
44 The Golden Rule of Accumulation

PART SEVEN – MACROECONOMIC POLICIES

45 Macroeconomic Policy
46 Monetary Policy
47 Classical, Keynesian and Modern Views on Monetary Policy
48 The Liquidity Theory of Money
49 Fiscal Policy
50 Monetarism versus Keynesianism
51 IS and LM Functions : General Equilibrium of Product and Money Market
52 Extensions of IS-LM Model
53 Effectiveness of Monetary and Fiscal Policy

PART EIGHT – MODERN MACROECONOMICS

54 The Rational Expectations Hypothesis
55 Supply-side Economics
56 The New Classical Macroeconomics
57 The Real Business Cycle Theory
58 New Keynesian Economics

PART NINE – MACROECONOMICS IN OPEN ECONOMY

59 Balance of Payments : Meaning and Components
60 Adjustment Mechanisms of Balance of Payments
61 Balance of Payments Policies : Internal and External Balance
62 Foreign Exchange Rate
63 Foreign Exchange Rate Policy

CONTENTS

Chapter-4
National Income Accounting

Chapter-5
The Circular Flow of Income

PART THREE – MACROECONOMIC THEORY

Chapter-6
The Classical Theory of Employment

Chapter-7
Say's Law of Market

Chapter-8
The Principle of Effective Demand : Aggregate Demand and Aggregate Supply

Chapter-9
The Consumption Function

Chapter-10
Theory of the Consumption Function

Chapter-11
The Investment Function

Chapter-12
The Concept of Multiplier

Chapter-13
Complex Multipliers

Chapter-14
Foreign Trade Multiplier

Chapter-15
The Principle of Acceleration and the Super Multiplier

Chapter-16
Some New Theories of Investment

Introduction
The Accelerator Theory of Investment
The Flexible Accelerator Theory or Lags in Investment
The Profits Theory of Investment
Duesenberry's Accelerator Theory of Investment
The Financial Theory of Investment
Jorgensons' Neoclassical Theory of Investment
Tobin's q Theory of Investment
Selected Readings
Exercises

Chapter-17
The Saving Function

Meaning of Saving Function
Determinants of Savings
The Paradox of Thrift
Exercises

Chapter-18
Saving and Investment Equality

The Classical View
The Keynesian View
Other Views
Selected Readings
Exercises

Chapter-19
The Model of National Income Determination

Introduction
Two-Sector Model
Three Sector Model
Four Sector Model : Income Determination In Open Economy
Exercises

Chapter-20
The Keynesian Theory of Income, Output and Employment : A Summary

Chapter-21
The Classical Vs. Keynesian Models of Income and Employment

Chapter-22
Unemployment and Full Employment

Chapter-23
Applicability of Keynes's Theory to Underdeveloped Countries

PART FOUR – MONETARY THEORY

Chapter-24
Money

Term Structure of Interest Rate

Meaning
Factors Determining the Term Structure of Interest Rates
Theories of Term Structure of Interest Rates
Exercises

Chapter-35
The Real Balance Effect and Pigou Effect

Patinkin's Integration of Monetary Theory and Value Theory : The Real Balance Effect
The Pigou Effect
Differences Between Pigou Effect and Real Balance Effect
Exercises

Chapter-36
Wage-Price Flexibility and Full Employment

Introduction
The Classical View
The Keynesian View
The Keynes Effect
Flexible Wage Policy vs Flexible Monetary Policy
Selected Readings
Exercises

PART FIVE – INFLATION AND BUSINESS CYCLES

Chapter-37
Inflation and Deflation

Introduction
Meaning of Inflation
The Inflationary Gap
Demand-Pull or Monetary Theories of Inflation
Cost-Push Inflation
Demand-Pull Versus Cost-Push Inflation
Mixed Demand-Pull Cost-Push Inflation
Sectoral or Demand-shift Inflation
Structural Inflation
Markup Inflation

Chapter-41
The Solow-Swan Model of Economic Growth

Chapter-42
The Endogenous Growth Theory

Chapter-43
Steady State Growth

Chapter-44
The Golden Rule of Accumulation

PART SEVEN – MACROECONOMIC POLICIES

Chapter-45
Macroeconomic Policy

Chapter-46
Monetary Policy

Chapter-47
Classical, Keynesian and Modern Views on Monetary Policy

Chapter-48
The Liquidity Theory of Money

Chapter-49
Fiscal Policy

Chapter-50
Monetarism versus Keynesianism

Chapter-51
IS and LM Functions : General Equilibrium of Product and Money Market

Chapter-52
Extensions of IS-LM Model

Chapter-53
Effectiveness of Monetary and Fiscal Policy

PART EIGHT – MODERN MACROECONOMICS

Chapter-54
The Rational Expectations Hypothesis

Introduction
Adaptive Expectations
Rational Expectations
Selected Readings
Exercises

Chapter-55
Supply-side Economics

Introduction
Main Features of Supply-side Economics
Policy Prescriptions of Supply-side Economics
Selected Readings
Exercises

Chapter-56
The New Classical Macroeconomics

Introduction
The New Classical Macroeconomics
The Labour Market
Policy Implications of New Classical Macroeconomics
Empirical Evidence
Selected Readings
Exercises

Chapter-57
The Real Business Cycle Theory

Introduction
Role of Technological Shocks
Criticisms of the Real Business Cycle Theory
Selected Readings
Exercises

Chapter-58
New Keynesian Economics

PART NINE – MACROECONOMICS IN OPEN ECONOMY

Chapter-59
Balance of Payments : Meaning and Components

Chapter-60
Adjustment Mechanisms of Balance of Payments

Chapter-61
Balance of Payments Policies : Internal and External Balance

Monetary-Fiscal Mix : Internal and External Balance Policies – Mundell-Fleming Model
Monetary and Fiscal Policies for Achieving Internal and External Balance Simultaneously – Swan Model
The Assignment Problem : The Mundellian Model of Monetary-Fiscal Policies for Internal and External Balance
Expenditure Switching Policies
Exercises

Chapter-62
Foreign Exchange Rate

Meaning of Foreign Exchange Rate
Determination of Equilibrium Exchange Rate
Theories of Foreign Exchange Rate
Causes of Changes in the Exchange Rate
Exercises

Chapter-63
Foreign Exchange Rate Policy

Introduction
Fixed Exchange Rates
Flexible Exchange Rates
Hybrid or Intermediate Exchange Rate Systems
Multiple Exchange Rates System
Exchange Rate Regimes in Practice
Exercises

PART-I

INTRODUCTION

CHAPTER 1

THE NATURE AND SCOPE OF MACROECONOMICS

INTRODUCTION

The term 'macro' was first used in economics by Ragner Frisch in 1933. But as a methodological approach to economic problems, it originated with the Mercantilists in the 16th and 17th centuries. They were concerned with the economic system as a whole. In the 18th century, the Physiocrats adopted it in their *Table Economique* to show the 'circulation of wealth' (*i.e.*, the net product) among the three classes represented by farmers, landowners and the sterile class. Malthus, Sismondi and Marx in the 19th century dealt with macroeconomic problems. Walras, Wicksell and Fisher were the modern contributors to the development of macroeconomic analysis before Keynes. Certain economists, like Cassel, Marshall, Pigou, Robertson, Hayek and Hawtrey, developed a theory of money and general prices in the decade following the First World War. But credit goes to Keynes who finally developed a general theory of income, output and employment in the wake of the Great Depression.

NATURE OF MACROECONOMICS

Macroeconomics[1] is the study of aggregates or averages covering the entire economy, such as total employement, national income, national output, total investment, total consumption, total savings, aggregate supply, aggregate demand, and general price level, wage level and cost structure. In other words, it is aggregative economics which examines the

interrelations among the various aggregates, their determination and causes of fluctuations in them. Thus in the words of Professor Ackley, "Macroeconomics deals with economic affairs 'in the large, it concerns the overall dimensions of economic life. It looks at the total size and shape and functioning of the "elephant" of economic experience, rather than working of articulation or dimensions of the individual parts. It studies the character of the forest, independently of the trees which compose it."[2]

Macroeconomics is also known as the theory of income and employment, or simply income analysis. It is concerned with the problems of unemployment, economic fluctuations, inflation or deflation, international trade and economic growth. It is the study of the causes of unemployment, and the various determinants of employment. In the field of business cycles, it concerns itself with the effect of investment on total output, total income, and aggregate employment. In the monetary sphere, it studies the effect of the total quantity of money on the general price level. In international trade, the problems of balance of payments and foreign aid fall within the purview of macroeconomic analysis. Above all, macroeconomic theory discusses the problems of determination of the total income of a country and causes of its fluctuations. Finally, it studies the factors that retard growth and those which bring the economy on the path of economic development.

The obverse of macroeconomics is microeconomics. Microeconomics is the study of the economic actions of individuals and small groups of individuals. The "study of particular firms, particular households, individual prices, wages, incomes, individual industries, particular commodities." But macroeconomics "deals with aggregates of these quantities; not with individual incomes but with the national income, not with individual prices but with the price levels, not with individual output but with the national output."[3] Microeconomics, according to Ackley, "deals with the *division* of total output among industries, products, and firms, and the *allocation of resources* among competing uses. It considers problems of income distribution. Its interest is in *relative* prices of particular goods and services." Macroeconomics, on the other hand, "concerns itself with such variables as the aggregate volume of the output of an economy, with the extent to which its resources are employed, with

the size of the national income, with the 'general price level'."[4]

Both microeconomics and macroeconomics involve the study of *aggregates*. But aggregation in microecnonmics is different from that in macroeconomics. In microeconomics the interrelationships of individual households, individual firms and individual industries to each other deal with aggregation. "The concept of 'industry', for example, aggregates numerous firms or even products. Consumer demand for shoes is an aggregate of the demands of many households, and the supply of shoes is an aggregate of the production of many firms. The demand and supply of labour in a locality are clearly aggregate concepts."[5] "However, the aggregates of microeconomic theory," according to Professor Bilas, "do not deal with the behaviour of the billions of dollars of consumer expenditures, business investments, and government expenditures. These are in the realm of microeconomics."[6] Thus the scope of microeconmics to aggregates relates to the economy as a whole, "together with sub-aggregates which (*a*) cross product and industry lines (such as the total production of consumer goods, or total production of capital goods), and which (*b*) add up to an aggregate for the whole economy (as total production of consumer goods and of capital goods add up to total production of the economy; or as total wage income and property income add up to national income)."[7] Thus microeconomics uses aggregates relating to individual households, firms and industries, while macroeconomics uses aggregates which relate them to the "economy wide total".

---

2. G. Ackley, *Macroeconomic Theory*, 1961, p.4.

3. K.E. Boulding, *Economic Analysis*, (3/e), p.237.

4. G. Ackley, *op. cit*., p.4.

Scope and Importance of Macroeconomics

As a method of economic analysis macroeconomics is of much theoretical and practical importance.

(1) To Understand the Working of the Economy. The study of

macroeconomic variables is indispensable for understanding the working of the economy. Our main economic problems are related to the behaviour of total income, output, employment and the general price level in the economy. These variables are statistically measurable, thereby facilitating the possibilities of analysing the effects on the functioning of the economy. As Tinbergen observes, macroeconomic concepts help in "making the elimination process understandable and transparent". For instance, one may not agree on the best method of measuring different prices, but the general price level is helpful in understanding the nature of the economy.

(2) In Economic Policies. Macroeconomics is extremely useful from the point of view of economic policy. Modern governments, especially of the underdeveloped economies, are confronted with innumerable national problems. They are the problems of overpopulation, inflation, balance of payments, general underproduction, etc. The main responsibility of these governments rests in the regulation and control of overpopulation, general prices, general volume of trade, general outputs, etc. Tinbergen says: "Working with macroeconomic concepts is a bare necessity *in order* to contribute to the solutions of the great problems of our times."[8] No government can solve these problems in terms of individual behaviour. Let us analyse the use of macroecomomic study in the solution of certain complex economic problems.

(i) In General Unemployment. The Keynesian theory of employment is an exercise in macroeconomics. The general level of employment in an economy depends upon effective demand which in turn depends on aggregate demand and aggregate supply fuctions. Unemployment is thus caused by deficiency of effective demand. In order to eliminate it, effective demand should be raised by increasing total investment, total output, total income and total consumption. Thus, macroeconomics has special significance in studying the causes, effects and remedies of general unemployment.

---

5. *Ibid*., pp. 4-5, n. 1.

6. R.A. Bilas, *Microeconomic Theory*, (1/e), p.1.

7. G. Ackley, *op. cit.*

8. J. Tinbergen, *op. cit*. Italics mine.

(ii) In National Income. The study of macroeconomics is very important for evaluating the overall performance of the economy in terms of national income. With the advent of the Great Depression of the 1930s, it became necessary to analyse the causes of general overproduction and general unemployment. This led to the construction of the data on national income. National income data help in forecasting the level of economic activity and to understand the distribution of income among different groups of people in the economy.

(iii) In Economic Growth. The economics of growth is also a study in macroeconomics. It is on the basis of macroeconomics that the resources and capabilities of an economy are evaluated. Plans for the overall increase in national income, output, employment are framed and implemented so as to raise the level of economic development of the economy as a whole.

(iv) In Monetary Problems. It is in terms of macroeconomics that monetary problems can be analysed and understood properly. Frequent changes in the value of money, inflation or deflation, affect the economy adversely. They can be counteracted by adopting monetary, fiscal and direct control measures for the economy as a whole.

(v) In Business Cycles. Further macroeconomics as an approach to economic problems started after the Great Depression. Thus its importance lies in analysing the causes of economic fluctuations and in providing remedies.

(3) For Understanding the Behaviour of Individual Units. For understanding the behaviour of individual units, the study of macroeconomics is imperative. Demand for individual products depends upon aggregate demand in the economy. Unless the causes of deficiency in aggregate demand are analysed, it is not possible to understand fully the reasons for a fall in the demand of individual products. The reasons for increase in costs of a particular firm or industry cannot be analysed without knowing the average cost conditions of the whole economy. Thus,

the study of individual units is not possible without macroeconomics.

Conclusion. We may conclude that macroeconomics enriches our knowledge of the functioning of an economy by studying the behaviour of national income, output, investment, saving and consumption. Moreover, it throws much light in solving the problems of unemployment, inflation, economic instability and economic growth.

Limitations of Macroeconomics

There are, however, certain limitations of macroeconomic analysis. Mostly, these stem from attempts to yield macroeconomic generalisations from individual experiences.

(1) Fallacy of Composition. In Macroeconomic analysis the "fallacy of composition" is involved, *i.e*., *aggregate economic behaviour is the sum total of individual activities*. But what is true of individuals is not necessarily true of the economy as a whole. For instance, savings are a private virtue but a public vice. If total savings in the economy increase, they may initiate a depression unless they are invested. Again, if an individual depositor withdraws his money from the bank there is no danger. But if all depositors do this simultaneously, there will be a run on the banks and the banking system will be adversely affected.

(2) To Regard the Aggregates as Homogeneous. The main defect in macro analysis is that it regards the aggregates as homogeneous without caring about their internal composition and structure. The average wage in a country is the sum total of wages in all occupations, *i.e*., wages of clerks, typists, teachers, nurses, etc. But the volume of aggregate employment depends on the relative structure of wages rather than on the average wage. If, for instance, wages of nurses increase but of typists fall, the average may remain unchanged. But if the employment of nurses falls a little and of typists rises much, aggregate employment would increase.

(3) Aggregate Variables may not be Important Necessarily. The aggregate variables which form the economic system may not be of much significance. For instance, the national income of a country is the total of all individual incomes. A rise in national income does not mean that

individual incomes have risen. The increase in national income might be the result of the increase in the incomes of a few rich people in the country. Thus a rise in the national income of this type has little significance from the point of view of the community.

Prof. Boulding calls these three difficulties as "macroeconomic paradoxes" which are true when applied to a single individual but which are untrue when applied to the economic system as a whole.[9]

(4) Indiscriminate Use of Macroeconomics Misleading. An indiscriminate and uncritical use of macroeconomics in analysing the problems of the real world can often be misleading. For instance, if the policy measures needed to achieve and maintain full employment in the economy are applied to structural unemployment in individual firms and industries, they become irrelevant. Similarly, measures aimed at controlling general prices cannot be applied with much advantage for controlling prices of individual products.

(5) Statistical and Conceptual Difficulties. The measurement of macroeconomic concepts involves a number of statistical and conceptual difficulties. These problems relate to the aggregation of microeconomic variables. If individual units are almost similar, aggregation does not present much difficulty. But if microeconomic variables relate to dissimilar individual units, their aggregation into one macroeconomic variable may be wrong and dangerous.

DIFFERENCE BETWEEN MICROECONOMICS AND MACROECONOMICS

The difference between microeconomics and macroeconomics can be made on the following counts.

The word *micro* has been derived from the Greek word *mikros* which means small. Microeconomics is the study of economic actions of individuals and small groups of individuals. It includes particular households, particular firms, particular industries, particular commodities and individual prices. Macroeconomics is also derived from the Greek word *makros* which means large. It "deals with aggregates of these

quantities, not with individual incomes but with the national income, not with individual prices but with the price levels, not with individual output but with the national output."

The objective of microeconomics on demand side is to maximize utility whereas on the supply side is to minimize profits at minimum cost. On the other hand, the main objectives of macroeconomics are full employment, price stability, economic growth and favourable balance of payments.

The basis of microeconomics is the price mechanism which operates with the help of demand and supply forces. These forces help to determine the equilibrium price in the market. On the other hand, the basis of macroeconomics is national income, output and employment which are determined by aggregate demand and aggregate supply.

---

9. K.E. Boulding, *A Reconstruction of Economics*, p. 173.

Microeconomics is based on different assumptions concerned with rational behaviour of individuals. Moreover the phrase *ceteris paribus* is used to explain the economic laws. On the other hand, macroeconomics bases its assumptions on such variables as the aggregate volume of output of an economy, with the extent to which its resources are employed, with the size of the national income and with the general price level.

Microeconomics is based on partial equilibrium analysis which helps to explain the equilibrium conditions of an individual, a firm, an industry and a factor. On the other hand, macroeconomics is based on general equilibrium analysis which is an extensive study of a number of economic variables, their interrelations and interdependences for understanding the working of the economic system as a whole.

In microeconomics, the study of equilibrium conditions are analysed at a particular period. But it does not explain the time element. Therefore, microeconomics is considered as a static analysis. On the other hand, macroeconomics is based on time-lags, rates of change, and past and expected values of the variables.

This rough division between micro and macroeconomics is not rigid, for the parts affect the whole and the whole affects the parts.

Dependence of Microeconomic Theory on Macroeconomics

Take for instance, when aggregate demand rises during a period of prosperity, the demand for individual products also rises. If this increase in demand is due to a reduction in the rate of interest, the demand for different types of capital goods will go up. This will lead to an increase in the demand for the particular types of labour needed for the capital goods industry. If the supply of such labour is less elastic, its wage rate will rise. The rise in wage rate is made possible by increase in profits as a consequence of increased demand for capital goods. Thus, a macroeconomic change brings about changes in the values of microeconomic variables—in the demands for particular goods, in the wage rates of particular industries, in the profits of particular firms and industries, and in the employment position of different groups of workers. Similarly, the overall size of income, output, employment, costs, etc. in the economy affects the composition of individual incomes, outputs, employment, and costs of individual firms and industries. To take another instance, when total output falls in a period of depression, the output of capital goods falls more than that of consumer goods. Profits, wages and employment decline more rapidly in capital goods industries than in the consumer goods industries.

Dependence of Macroeconomics on Microeconomic Theory

On the other hand, macroeconomic theory is also dependent on microeconomic analysis. The total is made up of the parts. National income is the sum of the incomes of individuals, households, firms and industries. Total savings, total investment and total consumption are the result of the saving, investment and consumption decisions of individual industries, firms, households and persons. The general price level is the average of all prices of individual goods and services. Similarly, the output of the economy is the sum of the output of all the individual producing units. Thus, "the aggregates and averages that are studied in

macroeconomics are nothing but aggregates and averages of the individual quantities which are studied in microeconomics."

---

10. K.E. Boulding, *Economic Analysis*, 3/e, p.237

Let us take a few concrete examples of this macro dependence on microeconomics. If the economy concentrates all its resources in producing only agricultural commodities, the total output of the economy will decline because the other sectors of the economy will be neglected. The total level of output, income and employment in the economy also depends upon income distribution. If there is unequal distribution of income so that income is concentrated in the hands of a few rich, it will tend to reduce the demand for consumer goods. Profits, investment and output will decline, unemployment will spread and ultimately the economy will be faced with depression. Thus, both macro and micro approaches to economic problems are interrelated and interdependent.

MACRO STATICS, MACRO DYNAMICS AND COMPARATIVE STATICS

Micro Statics

The word 'statics' is derived from the Greek word *statike* which means bringing to a standstill. In physics, it means a state of rest where there is no movement. In economics, it implies a state characterised by movement at a particular level without any change. It is a state, according to Clark, where five kinds of changes are conspicuous by their absence. The size of population, the supply of capital, methods of production, forms of business organisation and wants of the people remain constant, but the economy continues to work at steady pace. "It is to this active but unchanging process," writes Marshall, "that the expression static economics should be applied." Static economy is thus a timeless economy where no changes occur and it is necessarily in equilibrium. Indices are adjusted instantaneously: current demand, output and prices of goods and services. As pointed out by Prof. Samuelson: "Economic statics concerns itself with the simultaneous and instantaneous or timeless determination of economic variables by mutually interdependent relations." There is neither past nor future in the static state. Hence, there is no element of uncertainty

in it. Prof. Kuznets, therefore, believes that "static economics deals with relations and processes on the assumption of uniformity and persistence of either the absolute or relative economic quantities involved."

Macro-statics analysis explains the static equilibrium position of the economy. This is best explained by Professor Kurihara in these words, "If the object is to show a 'still picture' of the economy as a whole, the macro-static method is the appropriate technique. For this technique is one of investigating the relations between macro-variables in the final position of equilibrium without reference to the process of adjustment implicit in that final position."[11] Such a final position of equilibrium may be shown by the equation

$$Y = C + I.$$

where Y is the total income, C is the total consumption expenditure and I, the total investment expenditure. It simply shows a timeless identity equation without any adjusting mechanism. This macro-static model is illustrated in Figure 1.

According to this static Keynesian model, the level of national income is determined by the interaction of aggregate supply function and the aggregate demand function, In the Figure, 45° line represents the aggregate supply function and $C + I$ line, the aggregate demand fuction, 45° line and $C + I$ curve intersect at point E, the point of effective demand which determines OY level of national income.

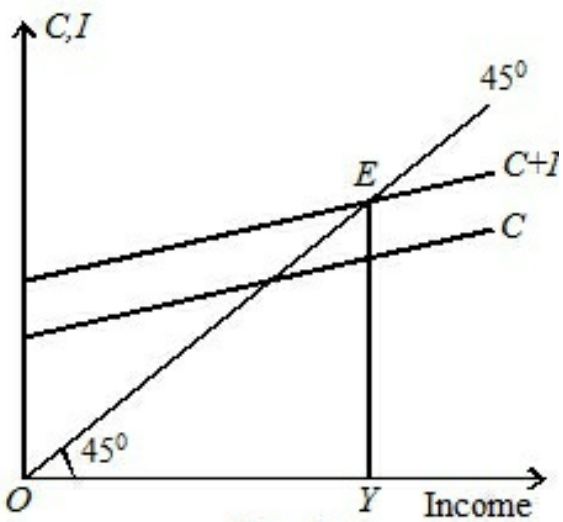
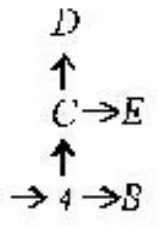


Fig. 1

Thus, economic statics refers to a timeless economy. It neither develops nor decays. It is like a snapshot photo from a 'still' camera which would be the same whether the previous and subsequent positions of the economy were subject to change or not.

Macro Dynamics



Economic dynamics, on the other hand, is the study of change, of acceleration or deceleration. It is the analysis of the process of change which continues *through time*. An economy may change through time in two ways: (a) without changing its pattern, and (b) by changing its pattern. Economic dynamics relates to the latter type of change. If there is a change in population, capital, techniques of production, forms of business organisation and tastes of the people, in any one or all of them, the economy will assume a different pattern, and the economic system will change its direction. In the accompanying diagram, given initial values of the economy, it would have proceeded along the path *AB*, but suddenly at *A* the indices change the pattern, and the direction of the equilibrium changes towards *C*. Again, it would have proceeded to *D* but at *C* the pattern and direction is changed to *E*. Thus, economic dynamics studies the *path* from one equilibrium position to another : from *A* to *C* and from *C* to *E*. Economic dynamics is, therefore, concerned with time-lags, rates of change, and past and expected values of the variables. In a dynamic economy, data change and the economic system takes time to adjust itself accordingly. According to Kurihara, "Macrodynamics treats discrete movements or rates of change of macro-variables. It enables one to see a 'motion-picture' of the functioning of the economy as a progressive whole."[12]

The macro-dynamic model is explained in terms of the Keynesian process of income propagation where consumption is a function of the income of the preceding period, *i.e* $C_t=f\,(Y_{t-1})$ and investment is a function of time and of constant autonomous investment ΔI,*i.e.*, I_t=f (Δ*I*). In Figure 2, *C* + *I* is the aggregate demand function and 45° line is the aggregate supply function. If we begin in period t_o where with an equilibrium level of income OY_0 , investment is increased by Δ*I*, then in period *t* income rises by the amount of the increased investment (from t_0 to *t*). The increased investment is shown by the new aggregate demand function *C*+*I*+Δ*I*. But in period *t*, consumption lags behind, and is still equal to the income at E_0. In period *t* + *I*, consumption rises and along with the new investment, it

increases income still higher to OY_1. This process of income propagation will continue till the aggregate demand function $C + I + \Delta I$ intersects the aggregate supply function 45° line at E_n in the nth period, and the new equilibrium level is determined at OY_n. The curved steps t_0 to E_n show the macro-dynamic equilibrium path.

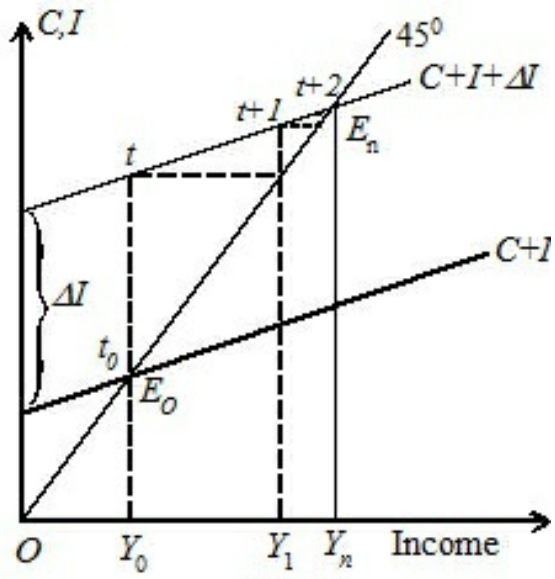


Fig. 2

11. K. Kurihara, *An Introduction to Keynesian Dynamics*, p.22

12. K.K. Kurihara, *op. cit.*, p.21.

Comparative Statics

Comparative statics is a method of economic analysis which was first used by the German economist, F.Oppenheimer in 1916. Schumpeter described it as "an evolutionary process by a succession of static models." In the words of Schumpeter, "Whenever we deal with disturbances of a given state by trying to indicate the static relations obtaining before a given disturbance impinged upon the system and after it had had time to work itself out. *This* method of procedure is known as Comparative Statics."[13] To be precise, comparative statics is the method of analysis in which different equilibrium situations are compared.

The distinction between static, comparative statics and dynamic situations is explained with the help of the accompanying figure. If the economy is working at situation *A* where it is producing at a *constant rate* without any change in the variables, it is a static state which is functioning *at a point* of time. When the economy moves from the equilibrium point *A* to point *B through time*, it is economic dynamics which traces out the actual path of movement of the economy between the two static equilibrium points. Comparative statics, on the other hand, is related to *once-over change* from point *A* to point *B* in which we do not study the forces behind the movement between the two points. Thus comparative statics is not

concerned with the transitional period but "involves the study of variations in equilibrium positions corresponding to specified changes in underlying data."

The Keynesian employment, income and output analysis is also based on the theory of shifting equilibrium wherein it $A \longleftrightarrow B$ compares different equilibrium levels of income. According to Kurihara, Keynes made no attempt to show the process of transition from one position of equilibrium to another. He simply used comparative statics analysis. Figure 3 explains two different levels of income, OY_2 at OT_1 time and OY_1 at OT_2 time. Independent of each other, both the income levels relate to economic statics. But income at OY_2 level is higher that at OY_1 level. This is comparative statics which compares two static levels of income as against dynamic economics which traces out the path *AB*, showing increase in income.

Limitations. But comparative statics is not without limitations.

1. Its scope is limited for it excludes many important economic problems. There are the problems of economic fluctuations and growth which can only be studied by the method of dynamic economics.

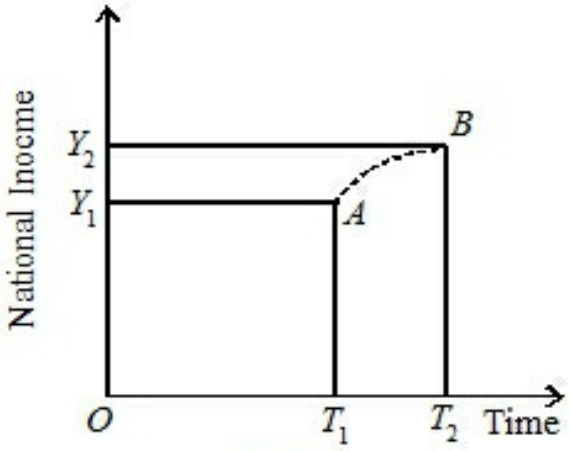


Fig. 3

2. Comparative statics is unable to explain the process of change from one position of equilibrium to another. It "gives only a partial glimpse of the movements, for we have only the two 'still pictures' to compare, whereas dynamics would give us a movie."

3. We are not sure when the new equilibrium will be established because this method neglects the transitional period. This makes comparative statics an incomplete and unrealistic method of economic analysis.

---

13. J.A. Schumpeter, *History of Economic Analysis*, 1954, p.965. Italics mine.

Conclusion

We sum up the discussion among macro statics, macro dynamics and comparative statics thus : Economic statics is the study of relations between economic variables *at a point of time*, whereas economic dynamics explains the relationship of economic variables *through time*. In a static economics there is movement but no change in economic phenomena while in dynamic economics, the fundamental forces themselves change. The former studies movement around the point of equilibrium, but the latter traces the path from one point of equilibrium, to the other, both backward or forward. On the other hand, comparative statics studies and compares two static equilibrium positions. If savings at a point of time are S_1 and at another moment of time S_2, this is once over change which is *comparative statics*. But if a given rise in savings leads to increase in investment, output, income and to a further rise in savings, this sequence of interdependent events of continuous changes is dynamic in nature.

No doubt economic dynamics is the antithesis of economic statics, yet the study of dynamic economics is a necessay adjunct to the hypothetical static analysis to enable economists to formulate generalisations. The *raison d'etre* of all static investigations is the explanation of dynamic change. On the other hand, dynamic economics is made up of static situations. If economic dynamics is the running picture of the working of the economy, economic statics relates to the 'still', the stationary position of the economy. Thus, both economic dynamics and economic statics are essential for the study and solution of economic problems.

TRANSITION FROM MICROECONOMICS TO MACROECONOMICS

As methodological approaches, both microeconomics and macroeconomics were used by the classical and neo-classical economists in their writings. But it was Marshall who developed and perfected microeconomics as a method of economic analysis. Similarly, it was Keynes who developed macroeconomics as a distinct method in economic theory. Therefore, the actual process of transition from microeconomics to macroeconomics started with the publication of Keynes's *General Theory*.

This transition has taken place in the following branches of economics.

Microeconomics is the study of economic actions of individuals and small groups of individuals. It includes particular households, particular firms, particular industries, particular commodities, individual prices, wages, and incomes. Thus microeconomics studies how resources are allocated to the production of particular goods and services and how efficiently they are distributed. But microeconomics, in itself, does not study the problem of allocation of resources to the economy as a whole. It is concerned with the study of parts and neglects the whole. As pointed out by Boulding, "Description of a large and complex universe of facts like the economic system is impossible in terms of individual items." Thus the study of microeconomics presents an imprecise picture of the ecomony. But the orthodox economists, like Pigou, tried to apply microeconomic analysis to the problems of an economy. Keynes thought otherwise and advocated macroeconomics which is the study of aggregates covering the entire economy such as total employment, total income, total output, total investment, total consumption, total savings, aggregate supply, aggregate demand, and general price level, wage level and cost structure. For understanding the problems facing the economy, Keynes adopted the macro approach and brought about the transition from micro to macro.

Microeconomics assumes the total volume of employment as given and studies how it is allocated among individual sectors of the economy. But Keynes rejected the assumption of full employment of resources, especially of labour. From the macro angle, he regarded full employment as a special case. The general situation is one of underemployment. The existence of involuntary unemployment of labour in capitalist economies proves that underemployment equilibrium is a normal situation and full employment is abnormal and accidental.

Keynes refuted Pigou's view that a cut in money wage could eliminate unemployment during a depression and bring about full employment in the economy. The fallacy in Pigou's agrument was that he extended the arguments to economy which were applicable to a particular industry. Reduction in money wage rate can increase employment in an industry by reducing its cost of production and the price of the product thereby raising

its demand. But the adoption of such a policy for the economy leads to a reduction in employment. When money wages of all workers in the economy are reduced, their incomes are reduced correspondingly. As a result, aggregate demand falls leading to a decline in employment in the economy as a whole.

Microeconomics takes the absolute price level as given and concerns itself with *relative prices* of goods and services. How the price of a particular commodity like rice, tea, milk, fan, scooter, etc. is determined ? How the wages of a particular type of labour, interest on a particular type of capital asset, rent on a particular land, and profits of an individual entrepreneur are determined ? But an economy is not concerned with relative prices but with the general level of prices. And the study of the *general level of prices* falls within the domain of macroeconomics. It is the rise or fall in the general price level that leads to inflation, and to prosperity and depression. Prior to the publication of Keynes's *General Theory*, economists concerned themselves with the determination of relative prices and failed to explain the causes of inflation and deflation or prosperity and depression. They attributed the rise or fall in the price level to the increase or decrease in the quantity of money. Keynes, on the other hand, showed that deflation and depression were caused by the deficiency of aggregate demand, and inflation and prosperity by the increase in aggregate demand. It is thus the rise or fall in aggregate demand which affects the general price level rather than the quantity of money.

Moreover, microeconomics being based on the assumption of full employment, it failed to provide an adequate explanation of the occurrence of trade cycles. It could not explain the turning points of the business cycles. By discarding the unrealistic assumption of full employment, Keynes and his followers have built models which not only explain the macroeconomic forces lying behind cyclical fluctuations but also explain the turning points of the cycle.

Another factor which has led to the transition from microeconomics to macroeconomics is the failure of microeconomics to deal with problems relating to the *growth* of the economy. Microeconomics concerns itself with the study of individual household, firm or industry. But principles

which are applicable to a particular household, firm or industry may not be applicable to the economy as a whole. This is because the level of aggregation differs in micro theory from macro theory. The classical economists committed the folly of applying micro theory to the economy as a whole while explaining economic growth. They emphasised the importance of saving or thrift in capital formation for economic growth. But in macro theory saving is a private virtue and a public vice. This is because increase in aggregate saving leads to a decline in aggregate consumption and demand, thereby decreasing the level of employent in the economy. Therefore, to remove unemployment and bring economic growth require increase in aggregate investment rather than saving. For economic growth, Harrod and Domar have emphasised the dual role of investment. *First*, it increases aggregate income, and *second*, it increases the productive capacity of the economy.

Microeconomics is based on the laissez-faire policy of a self-adjusting economic system with no government intervention. The classical economists were the votaries of *laissez-faire* policy. They believed in the automatic adjustment in the malfunctioning of the economy. They, therefore, had no faith either in monetary policy or fiscal policy for removing distortions in the economy. They also believed in the policy of balanced budgets. Keynes, who brought about the transition from micro to macro thinking, discarded the policy of laissez-faire. He believed that such a policy did not operate in public interest and it was this policy which had led to the Great Depression of 1930s. He, therefore, favoured state intervention and stressed the importance of deficit budgets during deflation and surplus budgets during inflation, along with cheap money and dear money policies respectively. The Keynesian policy measures have been adopted along with direct controls by the capitalist countries of the world.

STOCK AND FLOW CONCEPTS

The aggregates of macroeconomics are of two kinds. Some are stocks, typically the stock of capital *K* which is a timeless concept. Even in period analysis, a stock must be specified at a particular moment. Other aggregates are flows such as income and output, consumption and

investment. A flow variable has the time dimension *t*, as per unit of time or per period.

Stock is the quantity of an economic variable relating to a *point of time*. For example, store of cloth in a shop at a point of time is stock. Flow is the quantity of an economic variable relating to a *period of time*. The monthly income and expenditure of an individual, receipt of yearly interest rate on various deposits in a bank, sale of a commodity in a month are some examples of flow.

The concepts of stock and flow are used in the analysis of both microeconomics and macroeconomics.

In Microeconomics

In price theory or microeconomics, the concepts of stock and flow are related to the demand for and supply of goods. The market demand and supply of goods at a point of time are expressed as stock. The stock-demand curve of a good slopes downward from left to right like an ordinary demand curve, which depends upon price. But the stock-supply curve of a good is parallel to the Y-axis because the total quantity of stock of a good is constant at a point of time. On the other hand, the flow-demand and supply curves are like the ordinary demand and supply curves which are influenced by current prices. But the price is neither a stock nor a flow variable because it does not need a time dimenstion. Nor is it a stock quantity. In fact, it is a ratio between the flow of cash and the flow of goods.

In Macroeconomics

The concepts of stock and flow are used more in macroeconomics or in the theory of income, output, and employment. Money is a stock whereas the spending of money is a flow. Wealth is a stock and income is a flow. Saving by a person within a month is a flow while the total saving on a day is a stock. The government debt is a stock but the government deficit is a flow. The lending by a bank is a flow and its outstanding loan is a stock.

Some macro variables like imports, exports, wages, income, tax payments, social security benefits and dividends are always flows. Such flows do not have direct stocks but they can affect other stocks indirectly, just as imports can affect the stock of capital goods.

A stock can change due to flows but the size of flows can be determined itself by changes in stock. This can be explained by the relation between stock of capital and flow of investment. The stock of capital can only increase with the increase in the flow of investment, or by the difference between the flow of production of new capital goods and consumption of capital goods. On the other hand, the flow of investment itself depends upon the size of capital stock. But the stocks can affect flows only if the time period is so long that the desired change in stock can be brought about. Thus, flows cannot be influenced by changes in stock in the short run.

Lastly, both the concepts of stock and flow variables are very important in modern theories of income, output, employment, interest rate, business cycles, etc.

SELECTED READINGS

K.E. Boulding, *Economic Analysis*, Part II, 1996. Ch.13

G. Ackley, *Macroeconomic Theory*, 1961, Ch.1.

W.H. Branson, *Macroeconomic Theory and Policy*, 1972, Ch.1.

Edward Shapiro, *Macroeconomic Analysis*, 4/e, 1978, Ch.1.

F.S. Brooman, *Macro-Economics*, 4/e, 1970, Ch.1.

EXERCISES

1. Differentiate between Macrostatics and Macrodynamics and show that it is the study of Macrodynamics which is more significant.
2. Discuss the nature of macrostatics. What is the essential feature of

an economic system operating under static conditions ?

3. Which aspects of the analysis of the economy-wide magnitudes are 'macro' and which aspects 'static'? Suitably illustrate your answer.
4. Bring out the relationship that exists between micro and macro economics. Also distinguish between macro static and macro dynamics.
5. Distinguish between micro- and macro-economics. Explain the transition from microeconomics to macroeconomics.
6. What do you mean by macroeconomics ? Discuss its scope.
7. Discuss the scope of macroeconomics. What are the limitations of macroeconomic analysis ?
8. Write short notes on : (1) Macro-statics, (2) Macro-dynamics.
9. Distinguish between stock and flow.
10. Explain how macroeconomics describes the economy in terms of macro-economic variables.

PART-II

NATIONAL INCOME

CHAPTER 2

NATIONAL INCOME : MEANING AND MEASUREMENT

INTRODUCTION

National income is an uncertain term which is used interchangeably with national dividend, national output and national expenditure. On this basis, national income has been defined in a number of ways. In common parlance, national income means the total value of goods and services produced annually in a country. In other words, the total amount of income accruing to a country from economic activities in a year's time, is known as national income. It includes payments made to all resources in the form of wages, interest, rent and profits.

DEFINITIONS OF NATIONAL INCOME

The definitions of national income can be grouped into two classes: *One*, the traditional definitions advanced by Marshall, Pigou and Fisher; and *two*, modern definitions.

The Marshallian Definition

According to Marshall: "The labour and capital of a country acting on its natural resources produce annually a certain net aggregate of commodities, material and immaterial including services of all kinds. ...This is the true net annual income or revenue of the country or national dividend."[1] In this definition, the word 'net' refers to deductions from the

gross national income in respect of depreciation and wearing out of machines. And to this, must be added income from abroad.

Its Defects. Though the definition advanced by Marshall is simple and comprehensive, yet it suffers from a number of limitations. *First*, in the present day world, so varied and numerous are the goods and services produced that it is very difficult to have a correct estimation of them. Consequently, the national income cannot be calculated correctly. *Second*, there always exists the fear of the mistake of double counting, and hence the national income cannot be correctly estimated. Double counting means that a particular commodity or service like raw material or labour, etc. might get included in the national income twice or more than twice. For example, a peasant sells wheat worth Rs.2000 to a flour mill which sells wheat flour to the wholesaler and the wholesaler sells it to the retailer who, in turn, sells it to the customers. If each time, this wheat or its flour is taken into consideration, it will work out to Rs.8000, whereas, in actuality, there is only an increase of Rs.2000 in the national income. *Third*, it is again not possible to have a correct estimation of national income because many of the commodities produced are not marketed and the producer either keeps the produce for self-consumption or exchanges it for other commodities. It generally happens in an agriculture-oriented country country like India. Thus the volume of national income is underestimated.

The Pigouvian Definition

A.C. Pigou has in his definition of national income included that income which can be measured in terms of money. In the words of Pigou, "National income is that part of objective income of the community, including of course income derived from abroad which can be measured in money."[2] This definition is better than the Marshallian definition. It has proved to be more practical also. While calculating the national income now-a-days, estimates are prepared in accordance with the two criteria laid down in this definition. *First*, avoiding double counting, the goods and services which can be measured in money are included in national income. *Second*, income received on account of investment in foreign countries is included in national income.

Its Defects. The Pigouvian definition is precise, simple and practical but it is not free from criticism. *First*, in the light of the definition put forth by Pigou, we have to unnecessarily differentiate between commodities which can and which cannot be exchanged for money. But, in actuality, there is no difference in the fundamental forms of such commodities, no mattter they can be exchanged for money. *Second*, according to this definition when only such commodities as can be exchanged for money are included in estimation of national income, the national income cannot be correctly measured. According to Pigou, a woman's services as a nurse would be included in national income but excluded when she worked in the home to look after her children because she did not receive any salary for it. Similarly, Pigou is of the view that if a man marries his lady secretary, the national income diminishes as he has no longer to pay for her services. Thus the Pigovian definition gives rise to a number of paradoxes. *Third*, the Pigovian definition is applicable only to the developed countries where goods and services are exchanged for money in the market. According to this definition, in the backward and underdeveloped countries of the world, where a major portion of the produce is simply bartered, correct estimate of national income will not be possible, because it will always work out less than the real level of income. Thus the definition advanced by Pigou has a limited scope.

---

1. A. Marshall, *Principles of Economics*, p. 434.

2. A.C. Pigou, *The Economics of Welfare*, p. 31

Fisher's Definition

Fisher adopted 'consumption' as the criterion of national income whereas Marshall and Pigou regarded it to be production. According to Fisher, "The National dividend or income consists solely of services as received by ultimate consumers, whether from their material or from the human environments. Thus, a piano, or an overcoat made for me this year is not a part of this year's income, but an addition to the capital. Only the services rendered to me during this year by these things are income."[3] Fisher's definition is considered to be better than that of Marshall or Pigou, because Fisher's definition provides an adequate concept of economic

welfare which is dependent on consumption and consumption represents our standard of living.

Its Defects. But from the practical point of view, this definition is less useful, because there are certain dififculties in measuring the goods and services in terms of money. *First*, it is more difficult to estimate the money value of net consumption than that of net production. In one country there are several individuals who consume a particular good and that too at different places and, therefore, it is very difficult to estimate their total consumption in terms of money. *Second*, certain consumption goods are durable and last for many years. If we consider the example of piano or overcoat, as given by Fisher, only the services rendered for use during one year by them will be included in income. If an overcoat costs Rs. 100 and lasts for ten years, Fisher will take into account only Rs. 100 as national income during one year, whereas Marshall and Pigou will include Rs. 100 in the national income for the year, when it is made. Besides, it cannot be said with certainty that the overcoat will last only for ten years. It may last longer or for a shorter period. *Third*, the durable goods generally keep changing hands leading to a change in their ownership and value too. It, therefore, becomes difficult to measure in money the service-value of these goods from the point of view of consumption. For instance, the owner of a Maruti car sells it at a price higher than its real price and the purchaser after using it for a number of years, further sells it at its actual price. Now the question is as to which of its price, whether actual or black market one, should we take into account, and afterwards when it is transferred from one person to another, which of its value according to its average age should be included in national income ?

But the definitions advanced by Marshall, Pigou and Fisher are not altogether flawless. However, the Marshallian and Pigovian definitions tell us of the reasons influencing economic welfare, whereas Fisher's definition helps us compare economic welfare in different years.

Modern Definitions

From the modern point of view, Simon Kuznets has defined national income as "the net output of commodities and services flowing during the

year from the country's productive system in the hands of the ultimate consumers." On the other hand, in one of the reports of United Nations, national income has been defined on the basis of the systems of estimating national income, as net national product, as addition to the shares of different factors, and as net national expenditure in a country in a year's time. In practice, while estimating national income, any of these three definitions may be adopted, because the same national income would be derived, if different items were correctly included in the estimate.

---

3. I. Fisher, *The Nature of Capital and Income*, p. 104.

CONCEPTS OF NATIONAL INCOME

There are a number of concepts pertaining to national income and methods of measurement relating to them.

(A) Gross Domestic Product (GDP)

GDP is the total value of goods and services produced within the country during a year. This is calculated at market prices and is known as *GDP* at market prices. Dernberg defines GDP at market price as "the market value of the output of final goods and services produced in the domestic territory of a country during an accounting year."

There are three different ways to measure *GDP* : Product Method, Income Method and Expenditure Method. These three methods of calculating *GDP* yield the same result because National Product = National Income = National Expenditure.

1. The Product Method. In this method, the value of all goods and services produced in different industries during the year is added up. This is also known as the value added method to *GDP* or *GDP* at factor cost by industry of origin. The following items are included in India in this : agriculture and allied services; mining; manufacturing, construction, electricity, gas and water supply; transport, communication and trade; banking and insurance, real estates and ownership of dwellings and business services; and public administration and defence and other

services (or government services). In other words, it is the sum of gross value added.*

2. The Income Method. The people of a country who produce *GDP* during a year receive incomes from their work. Thus *GDP* by income method is the sum of all factor incomes : Wages and Salaries (compensation of employees) + Rent + Interest + Profit.

3. Expenditure Method. This method focuses on goods and services produced within the country during one year. *GDP* by expenditure method includes : (1) consumer expenditure on services and durable and non-durable goods (C),

(2) investment in fixed capital such as residential and non-residential building, machinery, and inventories (I),

(3) government expenditure on final goods and servies (G),

(4) export of goods and services produced by the people of country (X),

(5) *less* imports (M). That part of consumption, investment and government expenditure which is spent on imports is subtracted from *GDP*. Similarly, any imported component, such as raw materials, which is used in the manfacture of export goods is also excluded.

Thus *GDP* by expenditure method at market prices = C+ I + G + (X - M), where (X-M) is net export which can be positive or negative.

---

* For a detailed study of this method refer to "Value Added Approach to GDP."

(B) GDP at Factor Cost

GDP at factor cost is the sum of net value added by all producers within the country. Since the net value added gets distributed as income to the owners of factors of production, *GDP* is the sum of domestic factor incomes and fixed capital consumption (or depreciation).

Thus *GDP* at Factor Cost = Net value added + Depreciation.

GDP at factor cost includes : (i) compensation of employees i.e., wages, salaries, etc. (ii) operating surplus which is the business profit of both incorporated and unincorporated firms. [Operating Surplus = Gross Value Added at Factor Cost—Compensation of Employees—Depreciation] (iii) Mixed Income of Self-employed.

Conceptually, *GDP* at factor cost and *GDP* at market price must be identical. This is because the factor cost (payments to factors) of producing goods must equal the final value of goods and services at market prices. However, the market value of goods and services is different from the earnings of the factors of production. In *GDP* at *market price* are included indirect taxes and are excluded subsidies by the government. Therefore, in order to arrive at *GDP* at *factor cost*, indirect taxes are subtracted and subsidies are added to *GDP* at market price.

Thus, *GDP* at Factor Cost = *GDP* at Market Price - Indirect Taxes + Subsidies.

(C) Net Domestic Product (NDP)

NDP is the value of *net* output of the economy during the year. Some of the country's capital equipment wears out or becomes obsolete each year during the production process. The value of this capital consumption* is some percentage of gross investment which is deducted from *GDP*. Thus Net Domestic Product = *GDP* at Factor Cost - Depreciation.

(D) Nominal and Real GDP

When *GDP* is measured on the basis of current price, it is called *GDP* at current prices or nominal *GDP*. On the other hand, when *GDP* is calculated on the basis of fixed prices in some year, it is called *GDP* at constant prices or real *GDP*.

Nominal *GDP* is the value of goods and services produced in a year and measured in terms of rupees (money) at current (market) prices. In comparing one year with another, we are faced with the problem that the rupee is not a stable measure of purchasing power. *GDP* may rise a great deal in a year, not because the economy has been growing rapidly but

because of rise in prices (or inflation). On the contrary, *GDP* may increase as a result of fall in prices in a year but actually it may be less as compared to the last year. In both cases, *GDP* does not show the real state of the economy. To rectify the underestimation and overestimation of *GDP*, we need a measure that adjusts for rising and falling prices. This can be done by measuring *GDP* at constant prices which is called *real GDP*. To find out the real *GDP*, a base year is chosen when the general price level is normal, i.e., it is neither too high nor too low. The prices are set to 100 (or 1) in the base year. Now the general price level of the year for which real *GDP* is to be calculated is related to the base year on the basis of the following formula which is called the *deflator index* :

$$\text{Real } GDP = \frac{\text{GDP for the Current Year}}{} \times \frac{\text{Base Year (=100)}}{\text{Current Year Index}}$$

Suppose 1990-91 is the base year and *GDP* for 1999-2000 is Rs. 6,00,000 crores and the price index for this year is 300.

Thus, Real *GDP* for 1999-2000 = Rs. 6,00,000 x 100/300 = Rs. 2,00,000 crores

---

* Depreciation is called capital consumption.

(E) GDP Deflator

GDP deflator is an index of price changes of goods and services included in *GDP*. It is a price index which is calculated by dividing the nominal *GDP* in a given year by the real *GDP* for the same year and multiplying it by 100. Thus,

$$GDP \text{ Deflator} = \frac{\text{Nominal (or Current Prices) } GDP}{\text{Real (or Constant Prices) } GDP} \times 100$$

$$\text{For example, GDP Deflator in 1997-98} = \frac{1426.7\text{th. crores}}{1049.2\text{th. crores at}} \times 100$$

$$= 135.9$$

at 1993-94 prices

It shows that at constant prices (1993-94), *GDP* in 1997-98 increased by 135.9% due to inflation (or rise in prices) from Rs. 1049.2 thousand crores in 1993-94 to Rs. 1426.7 thousand crores in 1997-98.

(F) Gross National Product (GNP)

GNP is the total measure of the flow of goods and services at market value resulting from current production during a year in a country, including net income from abroad. *GNP* includes four types of final goods and services: (1) consumers' goods and services to satisfy the immediate wants of the people; (2) gross private domestic investment in capital goods consisting of fixed capital formation, residential construction and inventories of finished and unfinished goods; (3) goods and services produced by the government; and (4) net exports of goods and services, i.e., the difference between value of exports and imports of goods and services, known as net income from abroad.

In this concept of *GNP*, there are certain factors that have to be taken into consideration :

First, *GNP* is the measure of money, in which all kinds of goods and services produced in a country during one year are measured in terms of money at current prices and then added together. But in this manner, due to an increase or decrease in the prices, the *GNP* shows a rise or decline, which may not be real. To guard against erring on this account, a particular year (say for instance 1990-91) when prices be normal, is taken as the base year and the *GNP* is adjusted in accordance with the index number for that year. This will be known as *GNP* at 1990-91 prices or at constant prices.

Second, in estimating *GNP* of the economy, the market price of only the *final products* should be taken into account. Many of the products pass through a number of stages before they are ultimately purchased by consumers. If those products were counted at every stage, they would be included many a time in the national product. Consequently, the *GNP* would increase too much. To avoid double counting, therefore, only the final products, and not the intermediary goods should be taken into account.

Third, goods and services rendered *free* of charge are not included in the *GNP*, because it is not possible to have a correct estimate of their market price. For example, the bringing up of a child by the mother, imparting instructions to his son by a teacher, recitals to his friends by a musician, etc.

Fourth, the transactions which do not arise from the produce of *current* year or which do not contribute in any way to production, are not included in the *GNP*. The sale and purchase of old goods, and of shares, bonds and assets of existing companies are not included in *GNP* because these do not make any addition to the national product, and the goods are simply transferred.

Fifth, the payments received under *social security*, *e.g.*, unemployment insurance allowance, old age pension, and interest on public loans are also not included in *GNP*, because the recipients do not provide any service in lieu of them. But the depreciation of machines, plants and other capital goods is not deducted from *GNP*.

Sixth, the *profits* earned or *losses* incurred on account of changes in capital assets as a result of fluctuations in market prices are not included in the *GNP* if they are not responsible for current production or economic activity. For example, if the price of a house or a piece of land increases due to inflation, the profit earned by selling it will not be a part of *GNP*. But if, during the current year, a portion of a house is constructed anew, the increase in the value of the house (after subtracting the cost of the newly constructed portion) will be included in the *GNP*. Similarly, variations in the value of assets, that can be ascertained beforehand and are insured against flood or fire, are not included in the *GNP*.

Last, the income earned through illegal activities is not included in the *GNP*. Although the goods sold in the black market are priced and fulfil the needs of the people, but as they are not useful from the social point of view, the income received from their sale and purchase is always excluded from the *GNP*. There are two main reasons for this. *One*, it is not known whether these things were produced during the current year or the preceding years. *Two*, many of these goods are foreign made and smuggled and hence not included in the *GNP*.

Three Approaches to GNP

After having studied the fundamental constituents of *GNP*, it is essential to know how it is estimated. Three approaches are employed for this purpose. *One*, the income method to GNP; *two*, the expenditure method to *GNP*; and *three*, the value added method to *GNP*. Since gross income equals gross expenditure, *GNP* estimated by all these methods would be the same with appropriate adjustments.

1. Income Method to GNP

The income method to *GNP* consists of the remuneration paid in terms of money to the factors of production annually in a country. Thus *GNP* is the sum total of the following items :

(i) Wages and salaries. Under this head are included all forms of wages and salaries earned through productive activities by workers and entrepreneurs. It includes all sums received or deposited during a year by way of all types of contributions like overtime, commission, provident fund, insurance. etc.

(ii) Rents. Total rent includes the rents of land, shop, house, factory, etc. and the estimated rents of all such assets as are used by the owners themselves.

(iii) Interest. Under interest comes the income by way of interest received by the individual of a country from different sources. To this is added, the estimated interest on that private capital which is invested and not borrowed by the businessman in his personal business. But the interest received on governmental loans has to be excluded, because it is a mere transfer of national income.

(iv) Dividends. Dividends earned by the shareholders from companies are included in the *GNP*.

(v) Undistributed corporate profits. Profits which are not distributed by companies and are retained by them are included in the *GNP*.

(vi) Mixed incomes. These include profits of unincorporated business, self-employed persons and partnerships. They form part of *GNP*.

(vii) Direct taxes. Taxes levied on individuals, corporations and other businesses are included in the *GNP*.

(viii) Indirect taxes. The governement levies a number of indirect taxes, like excise duties and sales tax. These taxes are included in the price of commodities. But revenue from these goes to the government treasury and not to the factors of production. Therefore, the income due to such taxes is added to the *GNP*.

(ix) Depreciation. Every corporation makes allowance for expenditure on wearing out and depreciation of machines, plants and other capital equipment. Since this sum also is not a part of the income received by the factors of production, it is, therefore, also included in the *GNP*.

(x) Net income earned from abroad. This is the difference between the value of exports of goods and services and the value of imports of goods and services. If this difference is positive, it is added to the *GNP* and if it is negative, it is deducted from the *GNP*.

Thus *GNP* according to the Income Method = Wages and Salaries + Rents + Interest + Dividends + Undistributed Corporate Profits + Mixed Income + Direct Taxes + Indirect Taxes + Depreciation + Net Income from abroad.

2. Expenditure Method to GNP

From the expenditure view point, *GNP* is the sum total of expenditure incurred on goods and services during one year in a country. It includes the following items.:

(i) Private consumption expenditure. It includes all types of expenditure on personal consumption by the individuals of a country. It comprises expenses on durable goods like watch, bicycle, radio, etc., expenditure on single-used consumers' goods like milk, bread, ghee, clothes, etc., as also the expenditure incurred on services of all kinds like fees for school,

doctor, lawyer and transport. All these are taken as final goods.

(ii) Gross domestic private investment. Under this comes the expenditure incurred by private enterprise on new investment and on replacement of old capital. It includes expenditure on house construction, factory-buildings, all types of machinery, plants and capital equipment. In particular, the increase or decrease in inventory is added to or substracted from it. The inventory includes produced but unsold manufactured and semi-manufactured goods during the year and the stocks of raw materials, which have to be accounted for in *GNP*. It does not take into account the financial exchange of shares and stocks because their sale and purchase is not real investment. But depreciation is added.

(iii) Net foreign investment. It means the difference between exports and imports or export surplus. Every country exports to or imports from certain foreign countries. The imported goods are not produced within the country and hence cannot be included in national income, but the exported goods are manufactured within the country. Therefore, the difference of value between exports (X) and imports (M), whether positive or negative, is included in the *GNP*.

(iv) Government expenditure on goods and services. The expenditure incurred by the government on goods and services is a part of the *GNP*. Central, state or local governments spend a lot on their employees, police and army. To run the offices, the governments have also to spend on contingencies which include paper, pen, pencil and various types of stationery, cloth, furniture, cars, etc. It also includes the expenditure on government enterprises. But expenditure on transfer payments is not added, because these payments are not made in exchange for goods and services produced during the current year.

Thus *GNP* according to the Expenditure Method=Private Consumption Expenditure (C) + Gross Domestic Private Investment (I) + Net Foreign Investment (X–M) + Government Expenditure on Goods and Services (G) = C+ I + (X–M) + G.

As already pointed out above, *GNP* estimated by either the income or the expenditure method would work out to be the same, if all the items are

correctly calculated.

3. Value Added Method to GNP

Another method of measuring *GNP* is by value added. In calculating *GNP*, the money value of final goods and services produced at current prices during a year is taken into account. This is one of the ways to avoid double counting. But it is difficult to distinguish properly between a final product and an intermediate product. For instance, raw materials, semi-finished products, fuels and services, etc. are sold as inputs by one industry to the other. They may be final goods for one industry and intermediate for others. So, to avoid duplication, the value of *intermediate products* used in manufacturing final products must be subtracted from the value of total output of each industry in the economy. Thus, the difference between the value of material outputs and inputs at each stage of production is called the value added. If all such differences are added up for all industries in the economy, we arrive at the *GNP* by value added. *GNP* by value added = Gross value added + net income from abroad. Its calculation is shown in Tables 1,2 and 3.

Table 1 is constructed on the supposition that the entire economy for purposes of total production consists of three sectors. They are agriculture, manufacturing, and others, consisting of the tertiary sector. Out of the value of total output of each sector is deducted the value of its intermediate purchases (or primary inputs) to arrive at the value added for the entire economy. Thus the value of total output of the entire economy as per Table 1, is Rs. 155 crores and the value of its primary inputs comes to Rs. 80 crores. Thus the *GDP* by value added is Rs. 75 crores (Rs. 155 *minus* Rs. 80 crores).

TABLE 1 : GDP BY VALUE ADDED

s*(Rs. cror*

Industry	Total Output	Intermediate Purchases	Value Added
(1)	(2)	(3)	(4) = (2-3)
1. Agriculture	30	10	20

2. Manufacturing	70	45	25
3. Others	55	25	30
Total	155	80	75

TABLE 2
VALUE ADDED AT FACTOR COST

(Rs. Crores)

1.	Market Value of output	155
2.	*Less:* cost of intermediate Goods	80
3.	Gross value added	75
4.	*Less:* depreciation	8
5.	Net value added or domestic product at market prices	67
6.	*Less:* indirect taxes	7
7.	Net value added at factor cost	60

The total value added equals the value of gross domestic product of the economy. Out of this value added, the major portion goes in the form wages and salaries, rent, interest and profits, a small portion goes to the government as indirect taxes and the remaining amount is meant for depreciation. This is shown in Table 3.

Thus we find that the total gross value added of an economy equals the value of its gross domestic product. If depreciation is deducted from the gross value added, we have *net* value added which comes to Rs. 67 crores (Rs. 75 *minus* Rs. 8 crores). This is nothing but net domestic product at market prices. Again, if indirect taxes (Rs. 7 crores) are deducted from the net domestic product of Rs. 67 crores, we get Rs. 60 crores as the net value added at factor cost which is equivalent to net domestic product at factor cost. This is illustrated in Table 2.

TABLE 3 : GROSS DOMESTIC PRODUCT

(Rs Crores)

1. Wages and salaries	45
2. Income from rent	3
3. Net interest	4
4. Profits of companies	8
Net Value Added or NDP	60
5. Indirect taxes	+ 7
6. Depreciation	+ 8
Gross Value Added or GDP	75
7. Net income from abroad	+ 5
Gross National Income	80

Net value added at factor cost is equal to the net domestic product at factor cost, as given by the total of items 1 to 4 of Table 2 (Rs. 45+3+4+8 crores=Rs. 60 crores).

By adding indirect taxes (Rs 7 crores) and depreciation (Rs 8 crores), we get gross value added or GDP which comes to Rs 75 crores. If we add net income received from abroad to the gross value added, this gives us gross national income. Suppose net income from abroad is Rs. 5 crores. Then the gross national income is Rs. 80 crores (Rs. 75 crores + Rs. 5 crores) as shown in Table 3.

Its Importance. The value added method for measuring national income is more realistic than the product and income methods because it avoids the problem of double counting by excluding the value of intermediate products. Thus this method establishes the importance of intermediate products in the national economy. *Second*, by studying the national income accounts relating to value added, the contribution of each production sector to the value of the *GNP* can be found out. For instance, it can tell us whether agriculture is contributing more, or the share of manufacturing is falling, or of the tertiary sector is increasing in the current year as compared to some previous years. *Third*, this method is highly useful because "it provides a means of checking the *GNP* estimates

obtained by summing the various types of commodity purchases."

Its Difficulties. However, difficulties arise in the calculation of value added in the case of certain public services like police, military, health, education, etc. which cannot be estimated accurately in money terms. Similarly, it is difficult to estimate the contribution made to value added by profits earned on irrigation and power projects.

(G) GNP at Market Prices

When we multiply the total output produced in one year by their market prices prevalent during that year in a country, we get the Gross National Product at market prices. Thus *GNP* at market prices means the gross value of final goods and services produced annually in a country *plus* net income from abroad. It includes the gross value of output of all items from (1) to (4) mentioned under *GNP*. *GNP* at Market Prices = *GDP* at Market Prices + Net Income from Abroad.

(H) GNP at Factor Cost

GNP at factor cost is the sum of the money value of the income produced by and accruing to the various factors of production in one year in a country. It includes all items mentioned above under income method to *GNP less* indirect taxes. *GNP* at market prices always includes indirect taxes levied by the government on goods which raise their prices. But *GNP* at factor cost is the income which the factors of production receive in return for their services alone. It is the cost of production. Thus *GNP at market prices is always higher than GNP at factor cost*. Therefore, in order to arrive at *GNP* at factor cost, we *deduct* indirect taxes from *GNP* at market prices. Again, it often happens that the cost of production of a commodity to the producer is higher than a price of a similar commodity in the market. In order to protect such producers, the government helps them by granting monetary help in the form of a *subsidy* equal to the difference between the market price and the cost of production of the commodity. As a result, the price of the commodity to the producer is reduced and equals the market price of similar commodity. For example if the market price of rice is Rs. 3 per kg but it costs the producers in certain areas Rs. 3.50. The government gives a subsidy of 50 paise per kg to them

in order to meet their cost of production. Thus in order to arrive at *GNP* at factor cost, subsidies are *added* to *GNP* at market prices.

GNP at Factor Cost = *GNP* at Market Prices – Indirect Taxes + Subsidies.

(I) Net National Product (NNP)

NNP includes the value of total output of consumption goods and investment goods. But the process of production uses up a certain amount of fixed capital. Some fixed equipment wears out, its other components are damaged or destroyed, and still others are rendered obsolete through technological changes. All this process is termed *depreciation or capital consumption allowance*. In order to arrive at *NNP*, we *deduct* depreciation from *GNP*. The word 'net' refers to the exclusion of that part of total output which represents depreciation. So *NNP* = *GNP*—Depreciation.

(J) NNP at Market Prices

Net National Product at market prices is the *net* value of final goods and services evaluated at market prices in the course of one year in a country. If we deduct depreciation from *GNP* at market prices, we get *NNP* at market prices. So *NNP* at Market Prices = *GNP* at Market Prices—Depreciation.

(K) NNP at Factor Cost

Net National Product at factor cost is the *net* output evaluated at factor prices. It includes income earned by factors of production through participation in the production process such as wages and salaries, rents, profits, etc. It is also called *National Income*. This measure differs from *NNP* at market prices in that indirect taxes are *deducted* and subsidies are *added* to NNP at market prices in order to arrive at *NNP* at factor cost.[4] Thus

NNP at Factor Cost = NNP at Market Prices - Indirect taxes+ Subsidies
= *GNP* at Market Prices - Depreciation - Indirect taxes + Subsidies.
= National Income.

Normally, *NNP at market prices is higher than NNP at factor cost because indirect taxes exceed government subsidies. However, NNP at market prices can be less than NNP at factor cost when government subsidies exceed indirect taxes.*

(L) Domestic Income

Income generated (or earned) by factors of production within the country from its own resources is called domestic income or domestic product. Domestic income includes : (*i*) Wages and salaries, (*ii*) rents, including imputed house rents, (*iii*) interest, (*iv*) dividends, (*v*) undistributed corporate profits, including surpluses of public undertakings, (*vi*) mixed incomes consisting of profits of unincorporated firms, self-employed persons, partnerships, etc., and (*vii*) direct taxes.

Since domestic income does not include income earned from abroad, it can also be shown as : Domestic Income = National Income—Net income earned from abroad. Thus the difference between domestic income and national income is the net income earned from abroad. If we add net income from abroad to domestic income, we get national income, *i.e*., National Income = Domestic Income + Net income earned from abroad. But the net national income earned from abroad may be positive or negative. *If exports exceed import, net income earned from abroad is positive*. In this case, *national income is greater than domestic income*. On the other hand, when *imports exceed exports, net income earned from abroad is negative and domestic income is greater than national income*.[5]

---

4. In the relation between *NNP* at Market Prices and at Factor Cost, these items should be explained in detail as given under *GNP* at factor Cost.

(M) Private Income

Private income is income obtained by private individuals from any source, productive or otherwise, and the retained income of corporations. It can be arrived at from *NNP* at Factor Cost by making certain additions and deductions. The *additions* include transfer payments such as pensions, unemployment allowances, sickness and other social security benefits,

gifts and remittances from abroad, windfall gains from lotteries or from horse racing, and interest on public debt. The *deductions* include income from government departments as well as surpluses from public undertakings, and employees' contribution to social security schemes like provident funds, life insurance, etc.

Thus Private Income = National Income (or *NNP* at Factor Cost) + Transfer Payments + Interest on Public Debt — Social Security — Profits and Surpluses of Public Undertakings.

(N) Personal Income

Personal income is the total income received by the individuals of a country from all sources *before payment of direct taxes* in one year. Personal income is never equal to the national income, because the former includes the transfer payments whereas they are not included in national income. Personal income is derived from national income by deducting undistributed corporate profits, profit taxes, and employees' contributions to social security schemes. These three components are excluded from national income because they do reach individuals. But business and government transfer payments, and transfer payments from abroad in the form of gifts and remittances, windfall gains, and interest on public debt which are a source of income for individuals are added to national income. Thus Personal Income = National Income – Undistributed Corporate Profits – Profit Taxes – Social Security Contribution+ Transfer Payments + Interest on Public Debt.

Personal income differs from *private income* in that it is less than the latter because it excludes undistributed corporate profits.

Thus Personal Income = Private Income – Undistributed Corporate Profits – Profit Taxes.

(O) Disposable Income

Disposable income or personal disposable income means the actual income which can be spent on consumption by individuals and families. The whole of the personal income cannot be spent on consumption,

because it is the income that accrues before direct taxes have actually been paid. Therefore, in order to obtain disposable income, direct taxes are *deducted* from personal income. Thus Disposable Income=Personal Income – Direct Taxes.

---

5. Domestic Income may be 'net' or gross'. If not particularly stated it refers to *Net Domestic Product at Factor Cost* and can be distinguished from *Net Domestic Product at Market Prices as* in (K) above.

But the whole of disposable income is not spent on consumption and a part of it is saved. Therefore, disposable income is divided into consumption expenditure and savings. Thus Disposable Income = Consumption Expenditure + Savings.

If disposable income is to be deduced from national income, we *deduct* indirect taxes *plus* subsidies, direct taxes on personal and on business, social security payments, undistributed corporate profits or business savings from it and *add* transfer payments and net income from abroad to it.

Thus Disposable Income = National Income – Business Savings – Indirect Taxes + Subsidies – Direct Taxes on Persons – Direct Taxes on Business – Social Security Payments + Transfer Payments + Net Income from abroad.

(P) Real Income

Real income is national income expressed in terms of a general level of prices of a particular year taken as base. National income is the value of goods and services produced as expressed in terms of money at current prices. But it does not indicate the real state of the economy. It is possible that the net national product of goods and services this year might have been less than that of the last year, but owing to an increase in prices, *NNP* might be higher this year. On the contrary, it is also possible that *NNP* might have increased but the price level might have fallen, as a result national income would appear to be less than that of the last year. In both the situations, the national income does not depict the real state of the country. To rectify such a mistake, the concept of real income has been

evolved.

In order to find out the real income of a country, a particular year is taken as the base year when the general price level is neither too high nor too low and the price level for that year is assumed to be 100. Now the general level of prices of the given year for which the national income (real) is to be determined is assessed in accordance with the prices of the base year. For this purpose the following formula is employed.

$$\text{Real } NNP = NNP \text{ for the Current Year} \times \frac{\text{Base Year Index } (=100)}{\text{Current Year Index}}$$

Suppose 1990-91 is the base year and the national income for 1999-2000 is Rs. 20,000 crores and the index number for this year is 250. Hence, Real National Income for 1999-2000 will be = $20000 \times \frac{100}{250} =$ Rs. 8000 crores. This is also known as national income at constant prices.

(Q) Per Capita Income

The average income of the people of a country in a particular year is called Per Capita Income for that year. This concept also refers to the measurement of income at current prices and at constant prices. For instance, in order to find out the per capita income for 2001, at current prices, the national income of a country is divided by the population of the country in that year

$$\text{Per Capita Income for 2001} = \frac{\text{National income for 2001}}{\text{Population in 2001}}$$

Similarly, for the purpose of arriving at the Real Per Capita Income, this very formula is used.

$$\text{Real Per Capita Income for 2001} = \frac{\text{Real national income for 2001}}{\text{Population in 2001}}$$

This concept enables us to know the average income and the standard of

living of the people. But it is not very reliable, because in every country due to unequal distribution of national income, a major portion of it goes to the richer sections of the society and thus income received by the common man is lower than the per capita income.

METHODS OF MEASURING NATIONAL INCOME

There are four methods of measuring national income. Which method is to be used depends on the availability of data in a country and the purpose in hand.

(1) Product Method. According to this method, the total value of final goods and services produced in a country during a year is calculated at market prices. To find out the *GNP*, the data of all productive activities, such as agricultural products, wood received from forests, minerals received from mines, commodities produced by industries, the contributions to production made by transport, communications, insurance companies, lawyers, doctors, teachers, etc. are collected and assessed at market prices. Only the final goods and services are included and the intermediary goods and services are left out.[6]

(2) Income Method. According to this method, the net income payments received by all citizens of a country in a particular year are added up, *i.e.*, net incomes that accrue to all factors of production by way of net rents, net wages, net interest and net profits are all added together but incomes received in the form of transfer payments are not included in it. The data pertaining to income are obtained from different sources, for instance, from income tax department in respect of high income groups and in case of workers from their wage bills.[7]

(3) Expenditure Method. According to this method, the total expenditure incurred by the society in a particular year is added together and includes personal consumption expenditure, net domestic investment, government expenditure on goods and services, and net foreign investment. This concept is based on the assumption that national income equals national expenditure.[8]

(4) Value Added Method. Another method of measuring national income is the value added by industries. The difference between the value of material outputs and inputs at each stage of production is the value added. If all such differences are added up for all industries in the economy, we arrive at the gross domestic product.[9]

---

6. For details refer to the first para of 'Gross National Product'.

7. For details read 'Income Method to *GNP'* and its items.

8. For details read 'Expenditure Method to *GNP'* and its items.

9. For details read *'GNP'* by Value Added.

DIFFICULTIES OR LIMITATIONS IN MEASURING NATIONAL INCOME

There are many conceptual and statistical problems involved in measuring national income by the income method, product method, and expenditure method. We discuss them separately in the light of the three methods.

(A) Problems in Income Method

The following problems arise in the computation of National Income by income method.

1. Owner-occupied Houses. A person who rents a house to another earns rental income, but if he occupies the house himself, will the services of the house-owner be included in national income. The services of the owner-occupied house are included in national income as if the owner sells to himself as a tenant its services. For the purpose of national income accounts, the amount of imputed rent is estimated as the sum for which the owner-occupied house could have been rented. The imputed net rent is calculated as that portion of the amount that would have accrued to the house-owner after deducing all expenses.

2. Self-employed Persons. Another problem arises with regard to the income of self-employed persons. In their case, it is very difficult to find out the different inputs provided by the owner himself. He might be

contributing his capital, land, labour and his abilities in the business. But it is not possible to estimate the value of each factor input to production. So he gets a mixed income consisting of interest, rent, wage and profits for his factor services. This is included in national income.

3. Goods meant for Self-consumption. In under-developed countries like India, farmers keep a large portion of food and other goods produced on the farm for self-consumption. The problem is whether that part of the produce which is not sold in the market can be included in national income or not. If the farmer were to sell his entire produce in the market, he will have to buy what he needs for self-consumption out of his money income. If, instead he keeps some produce for his self-consumption, it has money value which must be included in national income.

4. Wages and Salaries paid in Kind. Another problem arises with regard to wages and salaries paid in kind to the employees in the form of free food, lodging, dress and other amenities. Payments in kind by employers are included in national income. This is because the employees would have received money income equal to the value of free food, lodging, etc. from the employer and spent the same in paying for food, lodging, etc.

(B) Problems in Product Method

The following problems arise in the computation of national income by product method :

1. Services of Housewives. The estimation of the unpaid services of the housewife in the national income presents a serious difficulty. A housewife renders a number of useful services like preparation of meals, serving, tailoring, mending, washing, cleaning, bringing up children, etc. She is not paid for them and her services are not include in national income. Such services performed by paid servants are included in national income. The national income is, therefore, underestimated by excluding the services of a housewife. The reason for the exclusion of her services from national income is that the love and affection of a housewife in performing her domestic work cannot be measured in monetary terms. That is why when the owner of a firm marries his lady secretary, her services are not included in national income when she stops working as a

secretary and becomes a housewife. When a teacher teaches his own children, his work is also not included in national income. Similarly, there are a number of goods and services which are difficult to be assessed in money terms for the reason stated above, such as painting, singing, dancing, etc. as hobbies.

2. Intermediate and Final Goods. The greatest difficulty in estimating national income by product method is the failure to distinguish properly between intermediate and final goods. There is always the possibility of including a good or service more than once, whereas only final goods are included in national income estimates. This leads to the problem of double counting which leads to the overestimation of national income.

3. Second-hand Goods and Assets. Another problem arises with regard to the sale and purchase of second-hand goods and assets. We find that old scooters, cars, houses, mechinery, etc. are transacted daily in the country. But they are not included in national income because they were counted in the national product in the year they were manufactured. If they are included every time they are bought and sold, national income would increase many times. Similarly, the sale and purchase of old stocks, shares, and bonds of companies are not included in national income because they were included in national income when the companies were started for the first time. Now they are simply financial transactions and represent claims. But the commission or fees charged by the brokers in the repurchase and resale of old shares, bonds, houses, cars or scooters, etc. are included in national income. For these are the payments they receive for their productive services during the year.

4. Illegal Activities. Income earned through illegal activities like gambling, smuggling, illicit extraction of wine, etc. is not included in national income. Such activities have value and satisfy the wants of the people but they are not considered productive from the point of view of society. But in countries like Nepal and Monaco where gambling is legalised, it is included in national income. Similarly, horse-racing is a legal activity in England and is included in national income.

5. Consumers' Service. There are a number of persons in society who render services to consumers but they do not produce anything tangible.

They are the actors, dancers, doctors, singers, teachers, musicians, lawyers, barbers, etc. The problem arises about the inclusion of their services in national income since they do not produce tangible commodities. But as they satisfy human wants and receive payments for their services, their services are included as final goods in estimating national income.

6. Capital Gains. The problem also arises with regard to capital gains. Capital gains arise when a capital asset such as a house, some other property, stocks or shares, etc. is sold at higher price than was paid for it at the time of purchase. Capital gains are excluded from national income because these do not arise from current economic activities. Similarly, capital losses are not taken into account while estimating national income.

7. Inventory Changes. All inventory changes (or changes in stocks) whether positive or negative are included in national income. The procedure is to take changes in physical units of inventories for the year valued at average current prices paid for them. The value of changes in inventories may be positive or negative which is added or subtracted from the current production of the firm. Remember, it is the change in inventories and not total inventories for the year that are taken into account in national income estimates.

8. Depreciation. Depreciation is deducted from *GNP* in order to arrive at *NNP*. Thus depreciation lowers the national income. But the problem is of estimating the current depreciated value of, say, a machine, whose expected life is supposed to be thirty years. Firms calculate the depreciation value on the original cost of machines for their expected life. This does not solve the problem because the prices of machines change almost every year.

9. Price Changes. National income by product method is measured by the value of final goods and services at current market prices. But prices do not remain stable. They rise or fall. When the price level rises, the national income also rises, though the national production might have fallen. On the contrary, with the fall in the price level, the national income also falls, though the national production might have increased. So price changes do not adequately measure national income. To solve this problem,

economists calculate the *real* national income at a constant price level by the consumer price index.

(C) Problems in Expenditure Method

The following problems arise in the calculation of national income by expenditure method :

(1) Government Services. In calculating national income by expenditure method, the problem of estimating government services arises. Government provides a number of services, such as police and military services, administrative and legal services. Should expenditure on government services be included in national income ? If they are final goods, then only they would be included in national income. On the other hand, if they are used as intermediate goods, meant for further production, they would not be included in national income. There are many divergent views on this issue. One view is that if police, military, legal and administrative services protect the lives, property and liberty of the people, they are treated as final goods and hence form part of national income. If they help in the smooth functioning of the production process by maintaining peace and security, then they are like intermediate goods that do not enter into national income, In reality, it is not possible to make a clear demarcation as to which service protects the people and which protects the productive process. Therefore, all such services are regarded as final goods and are included in national income.

(2) Transfer Payments. There arises the problem of including transfer payments in national income. Government makes payments in the form of pensions, unemployment allowance, subsidies, interest on national debt, etc. These are government expenditures but they are not included in national income because they are paid without adding anything to the production process during the current year. For instance, pensions and unemployment allowances are paid to individuals by the government without doing any productive work during the year. Subsidies tend to lower the market price of the commodities. Interest on national or public debt is also considered a transfer payment because it is paid by the government to individuals and firms on their past savings without any productive work.

(3) Durable-use Consummers' Goods. Durable-use consumers' goods also pose a problem. Such durable-use consumers' goods as scooters, cars, fans, TVs, furnitures, etc. are bought in one year but they are used for a number of years. Should they be included under investment expenditure or consumption expenditure in national income estimates ? The expenditure on them is regarded as final consumption expenditure because it is not possible to measure their used up value for the subsequent years.

But there is one exception. The expenditure on a *new house* is regarded as investment expenditure and not consumption expenditure. This is because the rental income or the imputed rent which the houseowner gets is for making investment on the new house. However, expenditure on a car by a household is consumption expenditure. But if he spends the amount for using it as a taxi, it is investment expenditure.

(4) Public Expenditure. Government spends on police, military, administrative and legal services, parks, street lighting, irrigation, museums, education, public health, roads, canals, buildings, etc. The problem is to find out which expenditure is consumption expenditure and which is investment expenditure. Expenses on education, museums, public health, police, parks, street lighting, civil and judicial administration are consumption expenditure. Expenses on roads, canals, buildings, etc. are investment expenditure. But expenses on defence equipment are treated as consumption expenditure because they are consumed during a war as they are destroyed or become obsolete. However, all such expenses including the salaries of armed personnel are included in national income.

IMPORTANCE OF NATIONAL INCOME ANALYSIS

The national income data have the following importance.

1. For the Economy. National income data are of great importance for the economy of a country. These days the national income data are regarded as accounts of the economy, which are known as *social accounts*. These refer to net national income and net national expenditure, which ultimately equal each other. Social accounts tell us how the aggregates of a nation's income, output and product result from the income of different

individuals, products of industries and transactions of international trade. Their main constituents are inter-related and each particular account can be used to verify the correctness of any other account.

2. National Policies. National income data form the basis of national policies such as employment policy, because these figures enable us to know the direction in which the industrial output, investment and savings, etc. change, and proper measures can be adopted to bring the economy to the right path.

3. Economic Planning. In the present age of planning, the national data are of great importance. For economic planning, it is essential that the data pertaining to a country's gross income, output, saving and consumption from different sources should be available. Without these, planning is not possible.

4. Economic Models. The economists propound short-run as well as long-run economic models or long-run investment models in which the national income data are very widely used.

5. Research. The national income data are also made use of by the research scholars of economics. They make use of the various data of the country's input, output, income, saving, consumption, investment, employment, etc., which are obtained from social accounts.

6. Per Capita Income. National income data are significant for a country's per capita income which reflects the economic welfare of the country. The higher the per capita income, the higher the economic welfare of the country.

7. Distribution of Income. National income statistics enable us to know about the distribution of income in the country. From the data pertaining to wages, rent, interest and profits, we learn of the disparities in the incomes of different sections of the society. Similarly, the regional distribution of income is revealed. It is only on the basis of these that the government can adopt measures to remove the inequalities in income distribution and to restore regional equilibrium. With a view to removing these personal and regional disequibria, the decisions to levy more taxes

and increase public expenditure also rest on national income statistics.

Inter-Relationships Among Different Concepts of National Income

The inter-relationships among the various concepts of national income can be shown in the form of equations as under :

1.	Gross National Product (GNP)	=	Gross National Expenditure (GNE)
2.	Gross Domestic Product (GDP)	=	GNP – Net Income from abroad.
3.	GNP at Market Prices	=	GNP at Factor Cost + Indirect Taxes – Subsidies
4.	NNP at Market Prices	=	GNP at Market Prices – Depreciation or Capital Consumption Allowance
5.	Net Domestic Product (NDP) at Market Prices	=	NNP at Market Prices – Net Factor Income from abroad
6.	NNP at Factor Cost or National Income or National Product	=	NNP at Market Prices – Indirect Taxes + Subsidies
7.	NDP at Factor Cost	=	National Income – Net Factor Income from abroad

	or Domestic Income or Domestic Product		
8.	Private Income	=	NNP at Factor Cost + Government and Business Transfer Payments + Current Transfers from abroad in the form of Gifts and Remittances + Windfall Gains + Net Factor Income from abroad + Interest on Public Debt and Consumer Interest – Social Security Contribution – Income from Government Departments and property – Profits and Surpluses of Public Corporations (or Undertakings) Or
		=	Income from Domestic Product accruing to Private Sector + Interest on Public Dept + Net Factor Income from abroad + Transfer Payments + Current Transfers from the rest of the world (or abroad)
9.	Income from Domestic Product accruing to Private Sector	=	NDP at Factor Cost – Income from Domestic Product accruing to Government Departments – Saving of Non-Departmental Enterprises.
10.	Personal Income	=	Private Income – Saving of Private Corporate Sector (or Undistributed Corporate Profits) – Corporation Tax (or Profit Taxes)
11.	Personal Disposable Income or Disposable Income	=	Personal Income – Direct Taxes paid by Households (or Direct Personal Taxes) and Miscellaneous Fees, Fines, etc. Or
		=	NDP at Factor Cost + Transfer Payments + Net

Factor Income from abroad – Corporation Tax – Undistributed Corporate Profits – Social Security Payments – Direct Personal Taxes

Or

= National Income at Factor Cost + Transfer Payments + Net Income from abroad – Corporate Tax – undistributed Corporate Profits – Social Security payments – Direct Personal Taxes – Indirect Taxes + Subsidies.

SOME SOLVED PROBLEMS

1. From the data pertaining to the Indian Economy given below, calculate (a) GNP at Factor Cost, (b) NNP at Factor Cost, (c) Net Domestic Product at Factor Cost, and (d) Net Domestic Product at market Prices.

		Rs. *Cro*
(i)	GNP at Market Prices	975
(ii)	Net factor income from abroad	(–) 2
(iii)	Capital consumption allowance	56
(iv)	Net indirect taxes	105

Solution

(a) GNP at Market Prices = 97503

GNP at Facor Cost = GNP at Market Prices – Indirect Taxes

= 97503 – 10576 = 86927

(b) NNP at Factor Cost = GNP at Factor Cost – Depreciation

= 86927 – 5699 = 81228

(c) NDP at Factor Cost = NNP at Factor Cost – Net Factor Income from abroad

(d) NDP at Market Prices = NDP at Factor Cost + Indirect Taxes

= 81429 + 10576 = 92005

2. On the basis of the following information relating to the Indian Economy, estimate (a) GNP at Market Prices, (b) Private Income, (c) Personal Income, and (d) Personal disposable Income.

		Rs. *Crores*
(i)	Net domestic product at factor cost	81429
(ii)	Income from domestic product accruing to the Government sector	2333
(iii)	Net factor income from abroad	(–) 201
(iv)	Indirect Taxes	12876
(v)	Subsidies	2300
(vi)	Interest on national debt	964
(vii)	Transfers from Government administrative departments	1981
(viii)	Other current transfers from the rest of the world	1271
(ix)	Corporation tax	1251
(x)	Undistributed corporate profits	464
(xi)	Direct taxes paid by households	2100
(xii)	Consumption of fixed capital	5699

Solution

(a) GNP at Market Prices = Net Domestic Product at Factor Cost + Net Factor Income from abroad + Capital Consumption Allowance (or Consumption of Fixed Capital) + Indirect Taxes – Subsidies

= 81429+(–201)+5699+12876–2300 = 97503

(b) Private Income = NNP at Factor Cost (or NI) + Transfer Payments + Interest on Public Dept – Social Security Contributions – Income from Government Departments – Profits and Surpluses of Public Corporations.

First, calculate NNP at Factor Cost because NDP at Factor Cost is given

in the illustration. So NNP at Factor Cost = NDP at Factor cost + Net Factor Income from abroad = 81429 + (–201) = 81228.

Now Private Income = 81228 + 1981 + 1271 + 964 – 2333 = 83111

where Transfer Payments from Government

Administrative Departments	=	1981
Other current transfers from the rest of the world	=	1271
Interest on National Debt	=	964
Income from Government Departments	=	2333

The remaining items in the equation are not given in our example.

Note : Items (vii) and (viii) relate to Transfer Payments.

(c) Personal Income	=	Private income – Undistributed Corporate Profit (or Savings of Private Corporate Sector) – Corporation Tax
	=	83111 – 464 – 1251
	=	81396
(d) Personal Disposable Income	=	Personal Income – Direct Taxes paid by Households
	=	81396 – 2100
	=	79296

3. On the basis of the following information, calculate Net National Product at Market Prices and Disposable Personal Income :

		Rs. *Crores*
(i)	Net domestic product at factor cost	15480
(ii)	Income from domestic product accruing to government	140
(iii)	National debt interest	170
(iv)	Transfer payments by government	240

(v) Net private donations from abroad	30
(vi) Net earned income from abroad	80
(vii) Indirect taxes	1330
(viii) Direct taxes	335
(ix) Subsidies	100
(x) Taxes on corporate profits	222
(xi) Undisributed profits of corporations	105

Solution

First calculate NNP at Factor Cost because NDP at Factor Cost is given in the illustration. So

NDP at Factor Cost	=	15480
+ Net earned income from abroad	=	80
∴ NNP at Factor Cost	=	15560

NNP at Market Prices

NNP at Factor Cost	=	15560
+ Indirect Taxes	=	1330
– Subsides	=	– 100
∴ NNP at Market Prices	=	16790

To Calculate Disposable Personal Income, calculate Private income and Personal Income.

Private Income

NNP at Factor Cost	=	15560
+ Transfer Payments by Government	=	240
+ Net private donations from abroad (These are transfers from the rest of the world)	=	30
+ National debt interest	=	170
– Income from domestic product accruing to	=	– 140

government	
∴ Private Income	15860

Personal Income

Private Income	=	15860
– Taxes on corporate profits	=	– 222
– Undistributed profits of corporations	=	– 105
∴ Personal Income	=	15533

Disposable Personal Income

Private Income	=	15533
– Direct Taxes on persons	=	–335
Disposable Personal Income	=	15198

EXERCISES

1. Explain the various concepts of national income. Under what circumstances national income tends to be underestimated ?
2. Point out the difference between the following concepts of National Income and also their relationship with each other :

   (a) Gross National Product and Net National Product.

   (b) National Income at Market Prices and National Income at Factor Costs.

   (c) Net National Product and Net Domestic Product.

   (d) Disposable Income and Personal Income.

3. Examine the usefulness of the concept of national income in analysing an economy's aggregate behaviour.

4. Explain the product and expenditure methods of computing national income. On what basis do they give identical results ?
5. What are the main conceptual problems involved in the estimation of *GNP* ? How can they be overcome ?
6. Explain the value added method of measuring national income.

CHAPTER 3

ECONOMIC WELFARE AND NATIONAL INCOME

WHAT IS ECONOMIC WELFARE ?

Before knowing the relation between economic welfare and national income, it is essential to define economic welfare. 'Welfare' is a state of the mind which reflects human happiness and satisfaction. In actuality, welfare is a happy state of human mind. Pigou regards individual welfare as the sum total of all satisfactions experienced by an individual and social welfare as the sum total of individual welfares. He divides welfare into economic welfare and non-economic welfare. Economic Welfare is that part of social welfare which can directly or indirectly be measured in money. Pigou attaches great importance to economic welfare because welfare is a very wide term. In his words : "The range of our enquiry becomes restricted to that part of social(general) welfare that can be brought directly or indirectly into relation with the measuring rod of money."[1] On the contrary, non-economic welfare is that part of social welfare which cannot be measured in money, for instance moral welfare.

But it is not proper to differentiate between economic and non-economic welfare on the basis of money. Pigou also accepts it. According to him, non-economic welfare can be improved upon in two ways. *First*, by the income-earning method. Longer hours of working and unfavourable conditions will affect economic welfare adversely. *Second*, by the income-spending method. It is assumed in economic welfare that expenditures incurred on different consumption goods provide the same amount of satisfaction. But in actuality it is not so, because when the utility of purchased goods starts diminishing the non-economic welfare declines

which results in reducing the total welfare. But Pigou is of the view that it is not possible to calculate such effects, because non-economic welfare cannot be measured in terms of money. The economist should, therefore, proceed with the assumption that the effect of economic causes on economic welfare applies also to total welfare. Hence, Pigou arrives at the conclusion that the increase in economic welfare results in the increase of total welfare and vice versa.

1. A.C. Pigou, *The Economics of Welfare*, 1932, p.11.

But it is not possible always, because the causes that lead to an increase in economic welfare may also reduce non-economic welfare. The increase in total welfare may, therefore, be less than anticipated. For instance, with the increase in income, both the economic welfare and total welfare increase and vice versa. But economic welfare depends not only on the amount of income but also on the methods of earning and spending it. When the workers earn more by working in factories but reside in slums and vitiated atmosphere, the total welfare cannot be said to have increased, even though the economic welfare might have increased. Similarly, as a result of increase in their expenditure proportionately to income, the total welfare cannot be presumed to have increased, if they spend their increased income on harmful commodities like wine, cigarettes etc. Hence, economic welfare is not an indicator of total welfare.

Relation between Economic Welfare and National Income

Pigou establishes a close relationship between economic welfare and national income, because both of them are measured in terms of money. When national income increases, total welfare also increases and vice versa. The effect of national income on economic welfare can be studied in two ways : (1) by change in the size of national income, and (2) by change in the distribution of national income.

Change in the Size of National Income

The change in the size of national income may be positive or negative.

The positive change in the national income increases its volume. As a result, people consume more of goods and services, which leads to increase in the economic welfare. Whereas the negative change in national income results in reduction of its volume. People get lesser goods and services for consumption which leads to decrease in economic welfare. But this relationship depends on a number of factors.

1. Change in Prices. Is the change in national income real or monetary ? If the change in national income is due to change in prices, it will be difficult to measure the real change in economic welfare. For example, when the national income increases as a result of increase in prices, the increase in economic welfare is not possible because it is probable that the output of goods and services may not have increased. It is more likely that the economic welfare would decline as a result of increase in prices. It is only the real increase in national income that increases economic welfare.

2. Working Conditions. It depends on the manner in which the increase in national income comes about. The economic welfare cannot be said to have increased, if the increase in national income is due to exploitation of labour e.g., increase in production by workers working for longer hours, by paying them lesser wages than the minimum. Forcing them to put their women and children to work, by not providing them with facilities of transport to and from the factories and of residence, and their residing in slums.

3. Per Capita Income. National income cannot be a reliable index of economic welfare, if per capita income is not kept in mind. It is possible that with the increase in national income, the population may increase at the same pace and thus the per capita income may not increase at all. In such a situation, the increase in national income will not result in increase in economic welfare. But from this, it should not be concluded that the increase in national income results in increase in economic welfare and vice versa.

It is possible that as a result of increase in national income, the per capita income might have risen. But if the national income has increased due to the production of capital goods and there is shortage of consumption goods on account of decrease in their output, the economic welfare will

not increase even if the national income and per capita income rise. This is because the economic welfare of people depends not on capital goods but on consumption goods used by them. Similarly, when the national income and the per capita income rise sharply during war time, the economic welfare does not increase because during war days the entire production capacity of the country is engaged in producing war material and there is shortage of consumption goods. As a result, the standard of living of the people falls and the economic welfare decreases.

Often, even with the increase in national income and per capita income the economic welfare decreases. This is the case when as a result of the increase in national income, income of the richer sections of the society increases and the poor do not gain at all from it. In other words, the rich become richer and the poor become poorer. Thus when the economic welfare of the rich increases and that of the poor decreases, the total economic welfare decreases.

4. Method of Spending. The influence of increase in national income on economic welfare depends also on the method of spending adopted by the people. If with the increase in income, people spend on such necessities and facilities as milk, ghee, eggs, fans, etc. which increase efficiency, the economic welfare will increase. But on the contrary, the expenditure on drinking, gambling etc. will result in decreasing the economic welfare. As a matter of fact, the increase or decrease in economic welfare as a result of increase in national income depends on changes in the tastes of people. If the change in fashions and tastes takes place in the direction of the consumption of better goods, the economic welfare increases.

Conclusion. It is clear from the above analysis that though the national income and economic welfare are closely inter-related, yet it cannot be said with certainty that the economic welfare would increase with the increase in national income and per capita income. The increase or decrease in economic welfare as a result of increase in national income depends on a number of factors such as the rate of growth of population, the methods of earning income, the conditions of working, the method of spending, the fashions and tastes, etc.

Changes in Distribution of National Income. The changes in the

distribution of national income takes place in two ways.

1. By Transfer of Wealth from the Poor to the Rich. When as a result of increase in national income, the transfer of wealth takes place in the former manner, the economic welfare decreases. This happens when the government gives more privileges to the richer sections and imposes regressive taxes on the poor.

2. Transfer of Wealth from the Rich to the Poor. The redistribution of wealth in favour of the poor is brought about by reducing the wealth of the rich and increasing the income of the poor. The income of the richer sections can be reduced by adopting a number of measures, e.g., by progressive taxation on income, property etc., by imposing checks on monopoly, by nationalising social services, by levying duties on costly and foreign goods which are used by the rich and so on. On the other hand, the income of the poor can also be raised in a number of ways, e.g., by fixing a minimum wage rate, by increasing the production of goods used by the poor, by fixing the prices of such goods, by granting financial assistance to the producers of these goods, by the distribution of goods through co-operative stores, and by providing free education, social security and low rent accommodation to the poor. When the distribution of income takes place in favour of the poor through these methods, the economic welfare increases.

But it is not essential that the equal distribution of national income would lead to increase in economic welfare. On the contrary, there is a greater possibility of economic welfare decreasing if the policy towards the rich is not rational. Heavy taxation and progressive taxes at high rates affect adversely the productive capacity, investment and capital formation, thereby decreasing the national income. Similarly, when through the efforts of the government, the income of the poor increases but if they spend that income on bad goods like drinking, gambling etc. or if their population increases, the economic welfare will decrease. Both these situations are not real and only express the fears, because the government, while imposing different kinds of progressive taxes on the rich, keeps particularly in view that taxation should not affect the production and investment adversely. On the other hand, when the income of a poor man

increases, he tries to provide better education to his children and to improve his standard of living, his welfare increases.

Conclusion. We arrive at the conclusion that as a result of the increase in national income, the economic welfare will increase provided that the income of the poor increases instead of decreasing and they improve their standard of living and that the income of the rich decreases in such a way that their productive capacity, investment and capital accumulation do not decline.

NATIONAL INCOME AS A MEASURE OF ECONOMIC WELFARE

GNP is not a satisfactory measure of economic welfare because the estimates of national income do not include certain services and production activities which affect welfare. We discuss below some of the factors which affect human welfare but are not included in the GNP estimates.

Leisure. One of the important things that affects the welfare of a society is leisure. But it is not included in GNP. For example, longer working hours may make people unhappy because their leisure is reduced. On the contrary, shorter working hours per week may increase leisure and make people happy. More or less leisure enjoyed by the community as such may affect the total output of the economy. But the value of leisure is excluded from the national income estimates.

Quality of Life. GNP estimates do not include the quality of life which reflects the community's welfare. Life in overcrowded cities is full of tensions. Roads are overcrowded. There is loss in time. Accidents occur daily which cripple or kill people. Environment becomes polluted. There are the problems of water, power, housing, transportation, etc. Crimes spread. Life becomes complex and the quality of life deteriorates. Consequently, social welfare is reduced. But all these stresses and strains of city life are not included in the national income estimates. Strangely, the efforts made by governments to remedy the ills of the city life are included in the GNP because they involve public expenditure.

On the other hand, in places where there is no congestion, people enjoy fresh air and the beauty of nature, the quality of life tends to increase. But this is not reflected in GNP.

Non-market Transactions. Some of the non-market transactions increase welfare but they are not included in national income estimates. The services of housewives within the home and community activities such as religious functions, affect the welfare of the people but they are excluded from the estimates of GNP because no market transaction is involved in providing these services.

Externalities. Similarly, there are externalities which tend to increase or decrease welfare but they are not included in GNP estimates. "An externality is a cost or benefit conferred upon second or third parties as a result of acts of individual production and consumption." But the cost or benefit of an externality cannot be measured in money terms because it is not included in market activities. An example of an external benefit is the pleasure one man derives from his neighbour's fine garden. An example of an external cost is environmental pollution caused by industrial plants. The former tends to increase welfare and the latter tends to reduce it. Since externalities are "untraded interdependencies", they are excluded from national income estimates.

Nature of Production. GNP estimates do not reflect the capacity of different goods to provide different levels of satisfaction to the community. The same amount of money spent on a nuclear bomb or on building a dam across a river adds equally to the national income. But they provide different levels of satisfaction to the community. A bomb does not increase welfare while a dam increases welfare.

Standard of Living. National income also does not reflect standard of living of the community which determines its welfare. If more national expenditure is incurred on the production of arms and ammunitions and on capital goods and less on producing consumption goods, this difference is not reflected in GNP estimates. But the reduction in the production of consumption goods tends to decrease the welfare of the people, while the increase in the expenditure on armaments and capital gooods does not increase welfare.

Keeping the above limitations in view, GNP cannot be used as a measure of welfare. However, a few economists have tried to broaden the definition of GNP so as to make it a measure of economic welfare. A pioneering attempt toward this direction has been made by Professors Nordhaus and Tobin[2] in 1972. They have constructed a 'Measure of Economic Welfare' which they call MEW. Professor Samuelson calls it 'Net Economic Welfare', or NEW.

According to Nordhaus and Tobin, in MEW they have tried to measure all consumption that leads to human welfare. To estimate the value of MEW, they deduct from consumption certain items which do not contribute to welfare and add other items that contribute to welfare but are excluded from GNP estimates.

The deductions which they make are of three types: (1) Those public and private expenditures that do not yield utility directly. They call them "regrettable necessities", such as government expenditures on national defence, police force, road maintenance, and sanitation services, and expenses by consumers on commuting (i.e., travelling regularly by train, scooter, car or bus between one's residence and place of work). (2) All consumer expenditures on durable household goods such as washing machines, cars, TV sets, etc. which yield utility over their lifetime. (3) Estimated costs arising from "negative externalities" which are disamenities arising from urbanisation, congestion and pollution. All these reduce human welfare.

---

2. William Nordhaus and James Tobin, 'Is growth obsolete,' in *Economic Growth*, National Bureau of Economic Research, New York, 1972

Having made these deductions, Nordhaus and Tobin add three items to consumption. They are: (1) the value of non-market activities; (2) the estimates of the value of the services of durable consumer goods actually consumed by the owners, both households and government; and (3) the estimates of the value of leisure.

In estimating MEW, Nordhaus and Tobin devote more attention to the valuation of leisure. For this they adopt two approaches: the opportunity

cost approach and the intrinsic-value approach. The *opportunity cost approach* is based on the principle that when a person chooses to enjoy more leisure, it is always at the cost of foregoing more income. An hour's leisure means an hour's wages foregone. They estimated that the value of leisure measured by the opportunity cost approach has been steadily rising over the years because of the steady rise in the real wage rate per hour over the years. The *intrinsic-value approach* measures the value of leisure in terms of the actual enjoyment (utility) provided by, say, an hour's leisure.

By using such valuation devices, Nordhaus and Tobin estimated that the figure of MEW in the United States for 1965 was dollars 1200 billion which was twice the GNP for the same year. Their estimate of the growth of per capita MEW for the period 1929-65 averaged 1.1 per cent a year, as against 1.7 per cent a year for per capita GNP for the same period. The estimates reveal that there was remarkable increase in economic welfare. At the same time, the regrettable necessities had also been growing rapidly.

From the above discussion, it should not be inferred that MEW is meant to replace GNP. It is at best an attempt to supplement GNP in order to include non-market activities in the latter for relating it to economic welfare.

SELECTED READINGS

A.C. Pigou, *The Economics of Welfare*, 4/e, 1932, Ch. VII.

P.A. Samuelson, *Economics*, 10/e, Ch. 10.

Edward Shapiro, *Macroeconomic Analysis*, 4/e, 1978, Appendix, pp. 62-65.

EXERCISE

1. How do changes in the size of national income and in the system of distribution affect economic welfare ? Explain.

2. What is economic welfare ? Clarify its relation with a country's national income.
3. Explain fully the effects of changes in the distribution of national dividend on economic welfare in the interest of the poor.
4. Is national income a satisfactory measure of economic welfare ? Give reasons. What alternative measure would you suggest?

CHAPTER 4

NATIONAL INCOME ACCOUNTING

SOCIAL ACCOUNTING

Meaning

The term 'social accounting' was first introduced into economics by J.R. Hicks in 1942. In his words, it means 'nothing else but the accounting of the whole community or nation, just as private accounting is the accounting of the individual firm'. Social accounting, also known as national income accounting, is a method to present statistically the inter-relationships between the different sectors of the economy for a thorough understanding of the economic conditions of the economy. It is a method of studying the *structure* of the body economic. It is a method of studying the structure of the body economic. It is a technique of presenting information about the nature of the economy with a view not merely to get an idea of its prosperity, past or present, but also to get guidelines for state policy to influence or regulate the economy. In the words of Edey, Peacock and Cooper : "Social accounting is concerned with the statistical classification of the activities of human beings and human institutions in ways which help us to understand the operation of the economy as a whole. The field of studies summed up by the words 'social accounting' embraces, however, not only the *classification* of economic activity, but also the *application* of the information thus assembled to the investigation of the operation of the economic system." In other words, social accounting describes statistically the economic activities of the different sectors of the entire economy, which indicates their mutual relationships and provides a framework for analysis.

COMPONENTS OF SOCIAL ACCOUNTING

The principal forms of ecnomic activity are production, consumption, capital accumulation, government transactions and transactions with the rest of the world. These are the components of social accounting. If the incomings and outgoings of a country relating to these five activities are shown in the form of accounts,they show a closed network of flows representing the basic structure of the economy. These flows are always expressed in money terms. We classify these flows as follows :

(1) Production Account. The production account relates to the business sector of the economy. It includes all forms of productive activity, i.e., manufacturing, trading, etc. It covers public and private companies, proprietory firms and partnerships, and state-owned business undertakings. Since all productive activity takes place within this sector, all payments flow from it to the other sectors. The production account of the business sector is shown in Table 1.

Table 1 : Production Account

(Rs Crores)

Payments		*Receipts*	
1.Payments to personal sector, i.e., wages, etc. (2-5)	279	5.Consumption expenditures (2-1)	219
2.Payments to government (3-5)	12	6. Government Purchases (3-1)	30
3.Business saving (4-3)	9	7.Gross private domestic investment (4-1)	36
4.Imports of goods and services (5-2)	9	8.Exports of goods and services (5-1)	24
Gross national income	309	Gross national expenditure	309

Note : Figures in brackets relate to corresponding Table and item number.

Payments to personal sector include rent, interest, dividend, wages, salaries, employees' compensation and proprietors' income. The item 'payments to government' includes producers' net payment to government in the form of taxes and social security payments. Business saving refers to producers' retained income or corporate saving. The last item relates to payments made to the foreign sector for imports of goods and services. These figures make up gross national income.

The receipt side of the production account shows the incomings to the business sector from sales of goods and services to the household or personal sector. Government purchases refer to goods and services sold by the business sector to the government. Gross private domestic investment comprises the gross flow of capital goods (fixed capital formation) and the net change in inventories. Net exports refer to the income earned by the business sector by selling goods and services to the rest of the world. The total of all these items gives GNP by expenditure.

(2) Consumption Account. The consumption account refers to the income and expenditure account of the *household* or *personal sector*. The household sector includes all consumers and non-profit making institutions such as clubs and associations. The consumption account is shown in Table 2.

Table 2 : Consumption Account

(Rs Crores)

Payments		*Receipts*	
1. Consumption expenditure (1-5)	219	5. Receipts from business, wages and salaries, etc. (1-1)	279
2. Payments to government (3-6)	45	6. Receipts from government (3-2)	6
3. Personal saving (4-4)	15		
4. Transfers to foreigners (5-3)	6	8.	
Personal outlay and saving	285	Personal income	285

Note : Figures in brackets relate to corresponding Table and item number.

The major item in the left side of the consumption account is the expenditure of household consumers in buying goods and services from the business sector to satisfy their wants. Payments to government include taxes and special insurance contributions. The next item refers to personal saving used for investment by the household sector. The item 'transfers to foreigners' might be taken to relate to investment in foreign securities or expenses by the residents on education or travel abroad. The right hand side of the account shows income of business and household consumers as the major item which comes in the form of wages and salaries, profit, interest, dividend, rent and receipts from current transfers, etc. Receipts from government include transfer paymants and net interest payments on public debt.

(3) Government Account. The government account relates to the outflows and inflows of the government sector. In the government sector are included all public authorities—centre, states and local authorities in a country. The government account is shown in Table 3.

Table 3 : Government Account

(Rs Crores)

Payments		*Receipts*	
1. Payments to business (1-6)	30	5. Receipts from business (1-2)	12
2. Payments to persons (2-6)	6	6. Receipts from persons (2-2)	45
3. Government surplus (4-5)	15		
4. Payments to foreigners (5-4)	6		
Government outlay and surplus	57	Government receipts	57

Note : Figures in brackets relate to corresponding Table and item number.

All items in the preceding Table have already been explained in the two accounts contrained in Tables 2 and 3, except item 3. This refers to investment made by the government out of its surplus or saving. However, the important point to be noted is that state-owned business enterprises are excluded from the government sector as they have been included in the business sector because like private enterprises public undertakings produce goods and services for sale.

(4) Capital Account. The capital account shows that saving equals domestic and foreign investment. Saving is invested in fixed capital and inventories within the country and/or in international assets. The capital account is shows in Table 4. The gross private investment includes the gross flow of capital goods and net change in inventories. Net foreign investment is the foreign surplus on current account. On the right side, gross saving includes business and personal savings and government surplus.

Table 4 : Capital Account

(Rs Crores)

Payments		*Receipts*	
1. Gross private domestic investment (1-7)	36	5. Business saving (1-3)	9
		4. Personal saving (2-3)	15
4. Net foreign investment (5-5)	3	5. Government surplus (3-3)	15
Gross investment	39	Gross saving	39

Note : Figures in brackets relate to corresponding Table and item number.

(5) Foreign Account. Foreign account shows the transactions of the country with the rest of the world. This account covers international movements of goods and services and transfer payments and corresponds to the current account of the international balance of payments. The foreign account or the rest-of-the world account is shown in Table 5. For simplicity, such services as freight and insurance have not been shown separately. All items have been already explained in the preceding accounts. It should be noted that in the foreign account 'exports' have been shown under payments (on the left side) and 'imports' under receipts (on the right side). This is because the amount received by the nationals of the country for exports is paid to foreign countries in exchange for imports and transfer payments. Here payments and receipts relate to the rest of the world and not to the country itself.

The five-account system detailed above relates to flows of the economy in terms of production, consumption, government transactions,

Table 5 : Foreign Account

(Rs Crores)

Payments		*Receipts*	
1.Exports of goods and services (1-8)	24	2Imports of goods and services (I-4)	9
		3.Transfer payments to foreigners by persons (2-4)	6
		4.Transfer payments to foreigners by government (3-4)	6
Net receipts from foreigners	24	5.Net foreign investment (4-2)	3
		Net payment to foreigners	24

Note : Figures in brackets relate to corresponding Table and item number.

capital accumulation, and transactions with the rest of the world. The accounts based on them are known as *functional accounts*, as they are based on a classification of transactions according to their functions.

Presentation of Social Accounts

Social accounts are presented on the double-entry basis like private accounts. Prevailing consensus is to present social accounts in the form of a social accounting table as recommended by the United Nations. A social accounting table is called a social accounts *matrix*. A transaction matrix is used for social accounts in which each row contains payments to other sectors and each column contains receipts form other sectors. Every single entry is both in a particular row and in a particular column. For balancing social accounts a row-total must equal its corresponding column-total. A matrix of social accounts is shown in Table 6 which presents the relationship between the flows of payments and receipts in accounts given in Tables 1 to 5.

Table 6 : Flow Matrix of Social Accounts

(Rs Crores)

Receipts from / Payments to	Accounts					
	1 Production	2 Consumption	3 Government	4 Capital	5 Foreign	Total
1. Production	-	279	12	9	9	309
2. Consumption	219	-	45	15	6	285
3. Government	30	6	-	15	6	57
4. Capital	36	-	-	-	3	39
5. Foreign	24	-	-	-	-	24
Total	309	285	57	39	24	714

In Table 6, each account has one row which shows the payments, and one column which shows the receipts, as explained below.

Row 1 shows payments made by the business sector to the tune of Rs. 279 crores to the consumption sector as wages, salaries, etc., Rs 12 crores to the government as taxes, Rs 9 crores as corporate saving to the capital account of firms and Rs 9 crores for importing goods and services from abroad.

Row 2 shows payments made to the business sector by the household sector amounting to Rs. 219 crores for buying goods and services from it, Rs 45 crores to the government in paying taxes, and insurance contributions, Rs. 15 crores to the investment (capital) sector in the form of saving by household consumers and Rs. 6 crores as investment in foreign securities, expenses on education, travel etc. in foreign countries.

Row 3 relates to the outflows of the government sector. The government pays Rs. 30 crores to the business sector for purchasing goods and services from it, Rs. 6 crores to the household sector as net interest payments on public debt and as transfer payments in the form of pension, gratuity, etc., Rs. 15 crores of government surplus is spent for investment purposes, and Rs. 6 crores are paid to foreign countries for goods and services received from them. The last item also includes expenditure on the maintenance of embassies abroad, and on delegations to foreign

countries.

Row 4 relates to the capital account of the economy out of which payment of Rs. 36 crores is made to business sector for capital goods and net change in investories, and Rs. 3 crores are net foreign investments.

Row 5 relates to the rest of the world account or foreign account to which payments of Rs. 24 crores are made by selling or exports of goods and services to foreigners.

Similarly, the receipts of each sector can be explained columnwise from Table 6.

The social accounts matrix presented in Table 6 further reveals three things. *First*, each cell (i.e., rectangular box) shows the equality of the payments to one sectoral account and the receipts from another sectoral account. For example, payment of Rs. 279 crores by the production sector to the consumption sector, reading row-wise in the Table is shown as the receipt of the consumption sector, reading column-wise. *Second*, the total payments of each sectoral account equal the total receipts of that sector. For example, the total payments of the production sector reading row-wise amounting to Rs. 309 crores equal the total receipts of this sector, reading column-wise. *Third*, the total payments of all sectors equal the total receipts of all sectors in the social accounting matrix. They are Rs. 714 crores both row-wise and column-wise in the Table.

Importance of Social Accounting

Social accounting helps in understanding the structure of an economy and relative importance of the different sectors and flows. It is a key to the evaluation and formulation of government policies both in the present and future.

The uses of social accounting are as follows :

(1) In Classifying Transactions. Economic activity in a country involves innumerable transactions relating to buying and selling, paying and receiving income, exporting and importing, paying taxes, etc. The great

merit of social accounting lies in classifying and summarising these different kinds of transactions properly, and deriving from these such aggregates as national income, national expenditure, saving, investment, consumption expenditure, production expenditure, government spending, foreign payments and receipts, etc.

(2) In Understanding Economic Structure. Social accounting helps us to understand the structure of the body economic. It tells us not only about the national income but also about the size of production and consumption, the level of taxation and saving and the dependence of the economy upon foreign trade.

(3) In Understanding Different Sectors and Flows. Social accounts throw light on the relative importance of the different sectors and flows in the economy. They tell us whether the contribution of the production sector, the consumption sector, the investment sector or the rest of the world sector is greater than the other sectors in the national accounts.

(4) In Clarifying Relations between Concepts. Social accounts help in clarifying the relationships between such related concepts as net national product at factor cost and gross national product at market prices.

(5) In Guiding the Investigator. Social accounts are a guide for the economic investigator by indicating the type of data which might be collected for analysing the behaviour of the economy. Such data might relate to gross national product, government expenditure on goods and services, private consumption expenditure, gross private investment, etc.

(6) In Explaining Trends in Income Distribution. Variations in the components of social accounts are a guide to the trends in income distribution within the economy.

(7) In Explaining Movements in GNP. Movements in gross national product valued at constant prices and expressed per head of population indicate changes in the standard of living. Similarly, changes in the level of productivity can be measured by relating gross national product valued at constant prices to working population per head.

(8) Provide a Picture of the Working of Economy. Social accounts provide an *ex post* picture of the working of the economy. "They can also be used as a framework for drawing up an *ex ante* forecast of the likely outcome of the economy in the future. Thus, social accounts ensure consistency of forecasts, both internally and in relation to other known facts."

(9) In Explaining Interdependence of Different Sectors of the Economy. Social accounts also provide an insight into the interdependence of the different sectors of the economy. This can be known from a study of the matrix of social accounts.

(10) In Estimating Effects of Government Policies. The importance of social accounts lies in estimating the effects of government policies on different sectors of the economy and in formulating new policies in keeping with changes in economic conditions, as revealed by national income accounts. Their main function is to help the government judge, guide or control economic conditions and to formulate economic policies which aim at maximisation of national income, keeping employment at a high level, reducing inequalities of income and wealth, preventing undue rise in prices, conserving foreign exchange, etc.

(11) Helpful in Big Business Organisations. Social accounts are also used by big business houses for assessing their performance and to improve their prospects on the basis of the statistical information about the various sectors of the economy.

(12) Useful for International Purposes. Social accounting is also useful for international purposes. A comparative study of the social accounts of different countries of the world helps in the categorization of countries into underdeveloped, less developed and developed. It is on the basis of social accounts that the various agencies of the United Nations make provisions for aid to poor countries of the world.

(13) Basis of Economic Models. Social accounts form the basis for economic models for the purpose of analysing the bahaviour of the economy as a whole, of economic forecasting and of illuminating problems of economic policy.

Difficulties of Social Accounting

The preparation of social accounts presents the following difficulties:

1. Imputations. In preparing social accounts, all incomes and payments are measured in money. But there are many goods and services which are difficult to impute in terms of money. They are services of the housewife in her home, painting as hobby by an individual, a teacher teaching his children at home, etc. Similarly there are a number of non-traded or non-marketed products and services. They are vegetables produced in the kitchen garden and consumed by the family itself, rental value of house occupied by the owner himself, a portion of farm produce retained by the farmer for personal consumption, etc. All such non-market transactions which cannot be assessed in money terms present problems in preparing social accounts accurately.

2. Double Counting. The greatest difficulty in preparing social accounts is of double counting. It arises from the failure to distinguish between final and intermediate products. For instance, flour used by a bakery is an intermediate product and that by a household the final product. Similarly, 'the purchase of a newly constructed building by the government is taken under *consumption output* of the economy. On the other hand, the purchase of the same building by a private firm becomes *gross investment* for the year'. Thus the same product is shown as consumption and investment in social accounts. Such problems lead to difficulties in preparing social accounts.

3. Public Services. Another problem is of estimating a number of public services in social accounts. They are police, military, health, education, etc. Similarly, the contributions made by multipurpose river valley projects cannot be fitted into the social accounts because of the difficulty of assessing their numerous benefits in monetary terms.

4. Inventory Adjustments. All inventory changes whether negative or positive are adjusted in the production accounts by inventory valuation adjustment. But the difficulty is that firms record inventories at their original costs and not at their replacement costs. When prices rise, there are gains in the book value of inventories. But when prices fall, there are

losses in the value of inventories, So for correct calculation of inventories in business accounts under social accounting, inventory valuation adjustment is required which is a very difficult thing.

5. Depreciation. Another problem in business accounts under social accounting is of estimating depreciation. For instance, it is very difficult to estimate the current depreciation rate of a capital asset whose expected life is very long, say fifty years. The difficulty increases further when prices of assets change every year. Unlike inventories, it is very difficult to have depreciation valuation adjustment in social accounts.

INPUT-OUTPUT ACCOUNTING

The input-output analysis tells us that there are industrial inter-relationships and inter-dependencies in the economic system as a whole. The inputs of one industry are the outputs of another industry and *vice versa*, so that ultimately their mutual relationships lead to equilibrium between supply and demand in the economy as a whole. Coal is an input for steel industry and steel is an input for coal industry, though both are the outputs of their respective industries. A major part of economic activity consists in producing intermediate goods (inputs) for further use in producing final goods (outputs). There are flows of goods in "whirlpools and cross currents" between different industries. The supply side consists of inter-industry flows of intermediate products and the demand side of the final goods. In essence, the input-output analysis implies that in equilibrium the money value of aggregate output of the whole economy must equal the sum of the money values of inter-industry inputs *plus* the sum of money values of inter-industry outputs.

The national income accounts are related to an economy's final product. They do not explicitly show the inter-industry flows of outputs and their relationships which the goods and services demanded. The input-output analysis analyses these relationships. It is, thus, an improvement over the national income accounting method.

Input-Output Table

The input-output accounting of national income is presented in an input-output table which is based on a 'transactions matrix'. A transactions matrix shows how the total output of one industry is distributed to all other industries as inputs and for final demand. A set of *m n* quantities or values arranged in *m* rows and *n* columns in a rectangular or square form is a matrix. That is why an input-output table is often called input-output matrix. The columns and rows of an input-output table 'provide industrial breakdowns of the final expenditures and income payments that enter into the national income accounts.'

A simple input-output matrix of an economy is shown in Table 7. Its rows show the amount of each industry's output sold to every other industry and to final buyers. The columns show the amount of each industry's inputs bought from every other industry, and from imports and factor services, known as primary inputs because they are not produced by the industries in the country.

Table 7 : Input-Output Transaction Matrix

(Rs Crores)

Purchasing Sectors → / Selling Sectors ↓	Inputs to			Final Demand (X+K+G+C)	Total Gross Output
	Agricul-ture	Manufac-turing	Others		
	1	*2*	*3*	*4*	*5*
Agriculture	-	15	5	22	42
Manufacturing	12	-	17	16	45
Others	8	12	-	30	50
Inports	7	5	8	7	27
Primary inputs	15	13	20	-	48
Total Gross Input	*42*	*45*	*50*	*75*	*212*

In this table, the total gross output of the agriculture sector of the economy is set in the first row (to be read horizontally). It consists of Rs. 15 crores to the manufacturing sector, Rs. 5 crores to the other sectors, and Rs. 22 crores to satisfy the final demand which comprises exports (X), capital (K), government (G) and personal consumption (C). Thus the total gross output of the agriculture sector is Rs. 42 crores = Rs. 20 crores of intermediate products (Rs. 15 crores *plus* Rs. 5 crores) + Rs. 22 crores of

final demand. Similarly, the second row shows the distributon of total output of the manufacturing sector of the economy valued at Rs. 45 crores per year. Likewise, the other rows show the distribution of output of other sectors, and from imports and primary inputs.

Taking columnwise (to be read downward), the first column shows inputs to the agriculture sector coming from the various sectors of the economy. For instance, inputs worth Rs. 12 crores come from the manufacturing industries, Rs. 8 crores from other sectors, Rs 7 crores from imports and Rs 15 crores from primary inputs. Primary inputs are the sum of payments as wages, profits, etc. and depreciation. They are also called *value added*. Thus the total gross input of the agriculture sector is 12+8+7+15=Rs. 42 crores. Similarly, the other columns show inputs to manufacturing and other sectors, and to final demand. The column relating to 'final demand' has been shown as nil against primary inputs. This means that the households of a country simply consume (or spend) but do not sell any thing to themselves. For instance, labour is not directly consumed.

It may be noted that the row total must equal the column total of the economy in the input-output table. It means that total gross output must equal the total gross input of the economy.

How to Find out GNP, GNI and GNE from the Input-Output Table?

Inter-industry transactions are not included in national income accounting. This is done in order to avoid the errors of multiple counting. In fact, intermediate goods (inputs and outputs) always enter into the production of goods. Thus only final demand or payments to factors enter into GNP at factor prices. In the preceding table, GNP *at factor prices* is Rs. 48 crores. The total resources available to the economy are GNP (primary inputs) *plus* imports: Rs. 48 crores + Rs 27 crores=Rs 75 crores. This is *Gross national Income* (GNI). GNI of Rs 75 crores is also the difference between total gross output and the total value of inputs or intermediate products, i.e., Rs 212 crores — Rs 137 crores = Rs 75 crores. Gross National Expenditure is the sum of payments to satisfy final demand which includes exports (X), capital expenditure (K), government expenditure (G) and consumption expenditure (C). Thus the total of final demand column in the table which is equal to Rs 75 crores

(=22+16+30+7) is the gross national expenditure (GNE) of the economy which equals GNI.

Input Co-efficient or Technical Co-efficient

There are two types of relationships which indicate and determine the manner in which an economy behaves and assumes a certain pattern of flows of resources. They are : (a) the internal stability or balance of each sector of the economy, and (b) the external stability of each sector or inter-sectoral relationships. Leontief calls them the "fundamental relationships of balance and structure." When expressed matematically, they are known as the "balance equations" and the "structural equations."

If the total output of say *Xi* of the *i*th industry be divided into various number of industries 1, 2, 3,...n, and the final demand *Di* then we have the balance equation :

$$Xi = xi_1 + xi_2 + xi_3 + ... xi_n + Di \quad ...(1)$$

and if the amount say *Yi* absorbed by the "outside sector" is also taken into consideration, then the balance equation of the *i*th industry becomes

$$Xi = xi_1 + xi_2 + xi_3 + ... xi_n + Di + Yi \quad ...(2)$$

It is to be noted that *Yi* stands for the sum of the flows of the products of the *i*th industry to consumption, investment and exports, net of imports, etc. It is also called the "final bill of goods" which is the function of the output to fill. Since xi_2 stands for the amount absorbed by industry 2 of the *i*th industry, it follows that *Xij* stands for the amount absorbed by the *j*th industry of *i*th industry.

The "technical co-efficient" or "input co-efficient" of the *j*th industry is denoted by:

$$aij = \frac{xij}{Xj}$$

Cross-multiplying, we have

$$xij = aij.Xj \qquad ...(3)$$

where *xij* is the flow from industry *i* to industry *j*, *Xj* is the total output of industry *j*; and *aij*, as already noted above, is a constant, called "technical co-efficient" or "flow" or "flow co-efficient" in the *i*th industry. Equation (3) is called a 'structural equation'.

The structural equation tells us that the output of one industry is absorbed by all indusries so that the flow-structure of the entire economy is revealed.

A number of structural equations *xij = aij.Xj* give a summary description of the economy's existing technological conditions. The table showing input co-efficients is called "a technology matrix".

The technology matrix of Table 7 is shown in Table 8.

These input co-efficients have been arrived at by dividing each item in the first column of Table 7 by its first row total and each item in the second column by its second row total and so on. Each column of the technological matrix reveals how much agriculture, manufacturing and other sectors require from each other to produce a rupee's worth of output. The first column shows that a rupee's worth of agriculture output requires inputs

Table 8. Technology Co-efficient Matrix (Input Co-efficient)

	Inputs to Agriculture	Inputs to Manufacturing	Inputs Oth
Agriculture	—	(15/45=) .33	(5/50=)
Manufacturing	(12/42=) .29	—	(17/50=)
Others	(9/42=) .19	(12/45=) .27	
(Primary inputs)	(22/45=) .52	(18/45=) .40	(28/50=)
Total	1.00	1.00	1

worth 29 paise from manufacturing, 19 paise from others and 52 paise from primary inputs.

The input co-efficient table can be utilised to measure the direct and indirect effects on the entire economy of any sectoral change in total output of final demand.

Limitations of Input-Output Accounting Analysis

Following are the limitations of input-output analysis:

1. Constancy of Input Coefficient Assumption Unrealistic. The input-output analysis has its shortcomings. Its framework rests on the assumption of constancy of input co-efficient of production. It tells us nothing as to how technical co-efficients would change with changed conditions. Again some industries may have identical capital structures, some may have heavy capital requirements while others may use no capital. Such variations in the use of techniques of production make the assumption of constant co-efficients of production unrealistic.

2. Factor Substitution Possible. This assumption of fixed co-efficients of production ignores the possibility of factor substitution. There is always the possibility of some substitutions even in a short period, while substitution possibilities are likely to be relatively greater over a longer period.

3. Rigid Model. The rigidity of the input-output model cannot reflect such phenomena as bottlenecks, increasing costs, etc.

4. Restrictive Model. The input-output model is severely simplified and restricted as it lays exclusive emphasis on the production side for the economy. It does not tell us why the inputs and outputs are of a particular pattern in the economy.

5. Difficulty in Final Demand. Another difficulty arises in the case of "final demand" or "bill of goods." In this analysis, the purchases by the government and consumers are taken as given and treated as a specific bill of goods. Final demand is regarded as an independent variable. It might,

therefore, fail to utilize all the factors proportionately or need more than their available supply. Assuming constancy of co-efficiency of production, the analysis is not in a position to solve this difficulty.

6. Quantity of Inputs not Constant. This analysis operates on the basis of a fixed quantity of an input for the production of per unit of output. As factors are mostly indivisible, the increases in outputs are not expected to be in proportion to the increases in inputs.

7. Solution of Equations Difficult. The input-output model works on equations which cannot be solved easily. First, the model of equations is prepared and then large numbers of data are collected. Equations require thorough knowledge of higher mathematics and even the collection of data is not so easy. This makes the construction of input-output model difficult.

Importance

Despite these limitations, the concept of input-output is of tremendous practical value and importance.

(1) A producer can know from the input-output table, the varieties and quantities of goods which he and the other firms buy and sell to each other. In this way, he can make the necessary adjustments and thus improve his position vis-a-vis other producers.

(2) It is also possible to find out from the input-output table the inter-relations among firms and industries about possible trends toward combinations.

(3) The effects of a prolonged strike, of a war and of a business cycle can be easily perceived from the input-output table.

(4) The input-output model has come to be used for national income accounting "because it provides a more detailed breakdown of the macro aggregates and money flows."

(5) The input-output analysis is also used for national economic planning.

The input-output model provides the necessary information about the structural co-efficients of the various sectors of the economy during a period of time or at a point of time which can be utilized for the optimum allocation of the economy's resources towards a desired end

FLOW OF FUNDS ACCOUNTS

The national income accounts do not tell anything about monetary or financial transactions whereby one sector places its savings at the disposal of the other sectors of the economy by means of loans, capital transfers, etc. In fact, the national income accounts do not take into consideration the financial dimensions of economic activity and they describe product accounts as if they are operated through barter. The flow of funds accounts are meant to supplement national income and product accounts. The flow of funds accounts were developed by Prof. Morris Copeland[1] in 1952 to overcome the weaknesses of national income accounting.

The flow of funds accounts list the sources of all funds received and the uses to which they are put within the economy. They show the financial transactions among different sectors of the economy and the link between saving and investment aggregates with lending and borrowing by them. The account for each sector reveals all the sources of funds whether from income or borrowing and all the uses to which they are put whether for spending or lending. This way of looking at financial transactions in their entirety has come to be known as the flow of funds approach or of sources and uses of funds.

In the flow of funds accounts, all changes in assets are recorded as uses and all changes in liabilities are recorded as sources. *Uses of funds* are increases in assets if positive or decreases in assets if negative. They refer to capital expenditures or *real investment* spending which involve the purchase of real assets. *Sources of funds* are increases in liabilities or *net worth* or *saving* if positive, and repayment of debt or dissaving if negative. Net worth is equal to a sector's total assets *minus* its total liabilities. Therefore a change in net worth equals any change in total assets *less* any change in total liabilities.

Flow of Funds Matrix

The flow of funds accounting system is presented in the form of a matrix by placing sources and uses of funds statements of different sectors side by side. It is an interlocking self-contained system that reveals financial relationships among all sectors of the economy. For the economy as a whole, total liabilities must equal total financial assets, although for any one sector its liabilities may not equal its financial assets. The consolidated net worth of an economy is consequently identical to the value of its real assets. This implies that saving must equal investment in an economy. Any single sector may save more than it invests or invest more than it saves. But the economy-wise total of saving must equal investment.

---

1. Morris A. Copeland, *A Study of Money Flows in the United States*, 1952

Table 9 presents the flow of funds matrix of an economy. For simplicity, we take the flow of funds accounts matrix of an economy divided into four sectors: households, nonfinancial corporations, financial institutions, and the government. These institutional sectors are shown in columns and various types of transactions in rows.

Table 9. Flow of Funds Accounts Matrix

(Rs crores)

Sectors / Transactions / Category		House-holds		Non-financial Corpora-tions		Financial Institu-tions		Govern-ment		Saving and invest-ment
		U	*S*	*U*	*S*	*U*	*S*	*U*	*S*	
1.	Gross saving	–	27	–	17	–	–	–	–4	40
2.	Gross investment	12	–	28	–	–	–	–	–	40
3.	Net financial invest-ment (4 – 5)	15	–	–11	–	–	–	–4	–	0
4.	Financial uses (net) (6+7+8+9+10)	25		3		6				34
5.	Financial sources (net) (6+7+8+9+10)		10		14		6		4	34
6.	Demand deposits	7	–	–1	–	–	6	–	–	0
7.	Government securities	4	–	2	–	–2	–	–	–	0
8.	Corporate securities	14	–	–	14	–	–	–	–	0
9.	Mortgages	–	10	–	–	8	–	2	–	0
10.	Net increase in FOREIGN ASSETS	–	–	2	–	+	–	–2	–	o

Note : 'U' refers to uses of funds and 'S' to sources of funds.

First take the columns. The household sector includes nonprofit organisations within it. Nonfinancial corporations include savings and loan associations, mutual savings banks, insurance companies, pension funds, mutual funds, etc. The remaining sectors are self-explanatory. The last column showing saving and investment is a measure of domestic saving and investment of all sectors *minus* the rest of the world.

Row 1 which relates to gross saving which is a *source* of funds for households (Rs 27 crores) and non-financial corporations (Rs 17 crores), and the minus figure of Rs. 4 crores for the government indicates a deficit in its budget.

Row 2 relates to gross investment which is a *use* of funds by households (Rs. 12 crores) and non-financial corporations (Rs 28 crores). The last column of the table shows that saving and investment are equal to Rs 40 crores each. The figures of saving and investment are supposed to have

been taken from the national income accounts of the economy.

Row 3 shows net financial investment which is the excess of saving over investment or uses over sources of each sector. For instance, the household sector makes positive net investment of Rs 15 crores (27-12), while the non-financial corporate sector incurs *negative* net investment of Rs 11 crores because it makes investment in excess of saving (17-28). The same is the case with the government which is shown as minus Rs 4 crores. (It can also be arrived at by deducting the figure of S of row 5 from the U figure of row 4 of each sector).

Row 4 shows financial uses (net) of funds. They feer to lending. It equals the sum of the change in each sector's holding of financial assets which include demand deposits, government securities, corporate securities, mortgages and net increase in foreign assets. Thus the net financial uses of the household sector are Rs 25 crores which include Rs 7 crores of demand deposits *plus* Rs 4 crores of government securities *plus* Rs 14 crores of corporate securities. Similarly for the remaining sectors.

Row 5 Financial sources (net) of funds shows the liability of each sector. They refer to borrowing. For instance, the government sector shows the acquisition of financial assets of Rs 4 crores by selling securities to the household sector.

Two important points sould be noted: *first*, financial uses (net) and financial sources (net) of the economy must equal. They are Rs 34 crores in our table. *Second*, changes in assets (uses) and liabilities (sources) of each type of fund must total up to zero. This is revealed by the last column of the table in relation to rows 6, 7, 8, 9 and 10. In the case of row 10 we have taken net increase in foreign assets to be zero for the sake of convenience. If it is a positive figure, the balance will show surplus in the international current account of the national income accounts and a negative figure will show a deficit.

Limitations

The flow of funds accounts are beset with a number of problems which are discussed as under:

1. The flow of funds accounts are more complicated than the national income accounts because they involve the aggregation of a large number of sectors with their very detailed financial transactions.

2. There is the problem of valuation of assets. Many assets, claims and obligations have no fixed value. It, therefore, becomes difficult to have their correct valuation.

3. The problem of inclusion of non-reproducible real assets arises in the flow of funds accounts. Economists have not been able to decide as to the type of reproducible assets which may be included in flow of funds accounts.

4. Similarly, economists have failed to decide about the inclusion of human wealth in flow of funds accounts.

Despite these problems, the flow of funds accounts supplement the national income accounts and help in understanding social accounts of an economy.

Importance

The flow of funds accounts present a comprehensive and systematic analysis of the financial transactions of the economy. As such, they are useful in a number of ways.

1. The flow of funds accounts are superior to the national income accounts. Even though the latter are fairly comprehensive, yet they do not reveal the financial transactions of the economy which the flow of funds accounts do.

2. They provide a useful framework for studying the behaviour of individual financial institutions of the economy.

3. According of Prof. Goldsmith, they bring "the various financial activities of an economy into explicit statistical relationships with one another and with data on the nonfinancial activities that generate income and production."

4. They trace the financial flows that interact with and influence the real saving-investment process. They record the various financial transactions underlying saving and investment.

5. They are essential raw materials for any comprehensive analysis of capital market behaviour. They help to identify the role of financial institutions in the generation of income, saving and expenditure, and the influence of economic activity on financial markets.

6. The flow of funds accounts show how the government finances its deficit and surplus budget and acquires financial assets.

7. They also show the results of transactions in government and corporate securities, net increase in deposits and foreign assets in the economy.

8. The flow of funds accounts help in analysing the impact of monetary policies on the economy as to whether they bring stability or instability or economic fluctuations.

DIFFERENCE BETWEEN FLOW OF FUNDS ACCOUNTS AND NATIONAL INCOME ACCOUNTS

The flow of funds accounts differ from national income accounts in many ways.

First, the national income accounts are confined exclusively to nonfinancial transactions. They neglect the link between saving and investment aggregates with lending and borrowing by different sectors of the economy.

Second, the national income accounts confine all real investment to the business sector with the exception of building construction. Consumers and governments are not allowed to invest in national income accounts. The flow of funds accounts treat consumer purchases of durable goods as real investment. Government enterprises are included in the producing sector of national income accounts but in the flow of funds accounts they are included in the government sector.

Third, the number of sectors in flow of funds accounts are more with larger details than in the national income accounts. They are defined institutionally in flow of funds accounts whereas they are defined functionally in national income accounts.

Fourth, there are fewer imputations in the flow of funds accounts than in national income accounts. For instance, taxes are carried on a cash basis in flow of funds accounts whereas some sectors are shown on an accrual basis in national income accounts.

BALANCE OF PAYMENTS ACCOUNTS

The balance of payments of a country is a systematic record of all its economic transactions with the outside world in a given year. It is a statistical record of the character and dimensions of the country's economic relationships with the rest of the world. According to Bo Sodersten, "The balance of payments is merely a way of listing receipts and payments in international transactions for a country." B.J. Cohen says, "It shows the country's trading position, changes in its net position as foreign lender or borower, and changes in its official reserve holding."

Structure and Classification

The balance of payments account of a country is constructed on the principle of double-entry book-keeping. Each transaction is entered on the credit and debit side of the balance sheet. But balance of payments accounting differs from business accounting in one respect. In business accounting, debits (–) are shown on the left side and credits (+) on the right side of the balance sheet. But in balance of payments accounting, the practice is to show credits on the left side and debits on the right side of the balance sheet.

When a payment is received from a foreign country, it is a credit transaction while payment to a foreign country is a debit transaction. The principal items shown on the *credit side* (+) are exports of goods and services, unrequited (or transfer) receipts in the form of gifts, grants etc. from foreigners, borrowings from abroad, investments by foreigners in the

country and official *sale* of reserve assets including gold to foreign countries and international agencies. The principal items on the *debit* side (–) include imports of goods and services, transfer (or unrequited) payments to foreigners as gifts, grants, etc., lending to foreign countries, investments by residents to foreign countries and official *purchase* of reserve assets or gold from foreign countries and international agencies.

These credit and debit items are shown vertically in the balance of payments account of a country according to the principle of double-entry book-keeping. Horizontally, they are divided into three categories : the current account, the capital account and the official settlements account or the official reserve assets account.

The balance of payments account of a country is constructed in Table 10.

Table 10. Balance of Payments Account

Credits (+) (Receipts)	*Debits (–) (Payments)*
1. Current Account	
Exports	*Imports*
(a) Goods	(a) Goods
(b) Services	(b) Services
(c) Transfer Payments	(c) Transfer Payments
2. Capital Account	
(a) Borrowings from Foreign Countries	(a) Lending to Foreign Countries
(b) Direct Investments by Foreign Countries	(b) Direct Investments in Foreign Countries
3. Official Settlements Account	
(a) Increase in Foreign Official Holdings	(a) Increase in Official Reserve of Gold and Foreign Currencies
Errors and Omissions	

1. Current Account. The current account of a country consists of all transactions relating to trade in goods and services and unilateral (or unrequited) transfers. Service transactions include costs of travel and transportation, insurance, income and payments of foreign investments, etc. Transfer payments relate to gifts, foreign aid, pensions, private remittances, charitable donations, etc. received from foreign individuals and governments to foreigners.

In the current account, merchandise exports and imports are the most important items. Exports are shown as a positive item and are calculated *f.o.b*. (free on board) which means that costs of transportation, insurance, etc. are excluded. On the other side, imports are shown as a negative item and are calculated *c.i.f.* (costs, insurance and freight) and included. The difference between exports and imports of a country is its *balance of visible trade* or merchandise trade or simply *balance of trade*. If visible exports exceed visible imports, the balance of trade is favourable. In the opposite case when imports exceed exports, it is unfavourable.

It is, however, services and transfer payments or invisible items of the current account that reflect the true picture of the balance of payments account. The balance of exports and imports of services and transfer payments is called the *balance of invisible trade*. The invisible items along with the visible items determine the actual current account position. If exports of goods and services exceed imports of goods and services, the balance of payments is said to be *favourable*. In the opposite case, it is unfavourable.

In the current account, the exports of goods and services and the receipts of transfer payments (unrequited receipts) are entered as credits (+) because they represent receipts from foreigners. On the other hand, the imports of goods and services and grant of transfer payments to foreigners are entered as debits (–) because they represent payments to foreigners. The net value of these visible and invisible trade balances is the balance on current account.

2. Capital Account. The capital account of a country consists of its transactions in financial assets in the form of short-term and long-term lendings and borrowings and private and official investments. In other

words, the capital account shows international flows of loans and investments, and represents a change in the country's foreign assets and liabilities. Long-term capital transactions relate to international capital movements with maturity of one year or more and include direct investments like building of a foreign plant, portfolio investment like the purchase of foreign bonds and stocks and international loans. On the other hand, short-term international capital transactions are for a period ranging between three months and less than one year.

There are two types of transactions in the capital account—private and government. Private transactions include all types of investment : direct, portfolio and short-term. Government transactions consist of loans to and from foreign official agencies.

In the capital account, borrowings from foreign countries and direct investment by foreign countries represent capital inflows. They are positive items or credits because these are receipts from foreigners. On the other hand, lending to foreign countries and direct investments in foreign countries represent capital outflows. They are negative items or debits because they are payments to foreigners. The net value of the balances of short-term and long-term direct and portfolio investments is the *balance on capital account*. The sum of current account and capital account is known as the *basic balance*.

3. The Official Settlements Account. The official settlements account or official reserve assets account is, in fact, a part of the capital account. But the U.K. and U.S. balance of payments accounts show it as a separate account. "The official settlements account measures the change in nations's liquidity and non-liquid liabilities to foreign official holders and the change in a nation's official reserve assets during the year. The official reserve assets of a country include its gold stock, holdings of its convertible foreign currencies and SDRs, and its net position in the IMF". It shows transactions in a country's net official reserve assets.

Errors and Omissions. Errors and omissions is a balancing item so that total credits and debits of the three accounts must equal in accordance with the principles of double entry book-keeping so that the balance of payments of a country always balances in the accounting sense.

IS BALANCE OF PAYMENTS ALWAYS IN EQUILIBRIUM?*

Balance of payments always balances means that the algebraic sum of the net credit and debit balances of current account, capital account and official settlements account must equal zero. Balance of payments is written as

$$B = R_f - P_f$$

where, B represents balance of payments,

R_f receipts from foreigners,

P_f payments made of foreigners

When $B = R_f - P_f = 0$, the balance of payments is in equilibrium.

When $R_f - P_f > 0$, it implies receipts from foreigners exceed payments made to foreigners and there is *surplus* in the balance of payments. On the other hand, when $R_f - P_f < 0$ or $R_f < P_f$ there is *deficit* in the balance of payments as the payments made to foreigners exceed receipts from foreigners.

If net foreign lending and investment abroad are taken, a flexible exchange rate creates an excess of exports over imports. The domestic currency depreciates in terms of other currencies. The exports becomes cheaper relatively to imports, It can be shown in equation form :

$$X + B = M + I_f$$

where X represents exports, M imports I_f foreign investment, B foreign borrowing

or $$X - M = I_f - B$$

or $$(X - M) - (I_f - B) = 0$$

The equation shows the balance of payments in equilibrium. Any positive balance in its current account is exactly offset by negative balance on its capital account and vice versa. In the accounting sense the balance of payments always balances. This can be shown with the help of the following equation :

$$C + S + T = C + I + G + (X - M)$$

or $$Y = C + I + G + (X - M) \quad (\therefore Y = C + S + T)$$

where C represents consumption exependiture, S domestic saving, T tax receipts, I investment expenditures, G government expenditures, X exports of goods and services and M imports of goods and services.

In the above equation

$C + S + T$ is GNI or national income (Y), and

$C + I + G = A$,

where A is called 'absorption'.

In the accounting sense, total domestic expenditures ($C + I + G$) must equal current income ($C + S + T$) that is $A = Y$. Moreover, domestic saving (S_d) must equal domestic investment (I_d). Similarly, an export surplus on current account ($X > M$) must be offset by an excess of domestic saving over investment ($S > I_d$). Thus the balance of payments always balances in the accounting sense, according to the basic principle of accounting. In the accounting system, the inflow and outflow of a transaction are recorded on the credit and debit sides respectively. Therefore, credit and debit sides always balance. If there is a deficit in the current account, it is offset by a matching surplus in the capital account by borrowings from abroad or/and withdrawing out of its gold and foreign exchange reserves, and vice versa. Thus, the balance of payments always balances in this sense also.

MEASURING DEFICIT OR SURPLUS IN BALANCE OF PAYMENTS

If the balance of payments always balances, then why does a deficit or surplus arise in the balance of payment of a country? It is only when all items in the balance of payments are included that there is no possibility of a deficit or surplus. But if some items are excluded from a country's balance of payments and then a balance is struck, it may show a deficit or surplus.

There are three ways of measuring deficit or surplus in the balance of payments.

First, there is the *basic balance* which includes the current account balance and the long-term capital account balance.

Second, there is the *net liquidity balance* which includes the basic balance and the short-term private non-liquid capital balance, allocation of SDRs, and errors and omissions.

Third, there is the *official settlements balance* which includes the total net liquid balance and short-term private liquid capital balance.

If the total debits are more than total credits in the current and capital accounts, including errors and omissions, the net *debit balance* measures the *deficit* in the balance of payments of a country. This deficit can be settled with an equal amount of *net credit balance* in the official settlements account. On the contrary, if total credits are more than total debits in the current and capital accounts, including errors and omissions, the *net debit* balance measures the *surplus* in the balance of payments of a country. This surplus can be settled with an equal amount of *net debit* balance in the official settlements account.

The relationship between these balances is summarised in Table 11.

Table 11

Trade balance.....*a*	
Transfer payment balance.....*b*	Autonomous
Current Account Balance.....*c* (= *a* + *b*)	Items
Long-term capital balance.....*d*	
Basic Balance.....*e* (= *c* + *d*)	
Short-term private non-liquid capital balance.....*f*	
Allocation of SDRs.....*g*	Accommodating
Errors and omissions.....*h*	Items
Net Liquidity Balance.....*i* (= *e* + *f* + *g* + *h*)	
Short-term private liquid capital balance.....*j*	
Official Settlements Balance.....*k* (= *i* + *j*)	

Each balance would give different figure of the deficit. The items that are included in a particular balance are placed 'above the line' and those excluded are put ' below the line'. Items that are put above the line are called *autonomous items*. Items that are placed below the line are called *settlement or accommodating or compensatory or induced items.* All transactions in the current and capital accounts are autonomous items because they are undertaken for business or profit motives and are independent of balance of paymentts considerations. According to Sodersten and Reed, "Transactions are said to be autonomous if their value is determined independently of the balance of payments". Whether there is BOP deficit or surplus depends on the balance of autonomous items. If autonomous receipts are less than autonomous payments, BOP is in deficit and vice versa.

"Accommodating items on the other hand are determined by the net consequences of the autonomous items", according to Sodersten and Reed. They are in the official reserve account. They are compensating (induced or accommodating) short-term capital transactions which are meant to correct a disequilibrium in the autonomous items of balance of payments.

But it is difficult to determine which item is compensatory and which is autonomous. For instance, in the table given above, the main difference in

the three balances is their treatment of short-term capital movements which are responsible for deficit in the balance of payments. The basic balance places short-term private non-liquid capital movements below the line while the net liquid balance puts them above the line. Similarly, the net liquid balance places short-term private liquid capital movements below the line and the official settlements balance puts them above the line. Thus, as pointed out by Sodersten and Reed, "Essentially the distinction between autonomous and accommodating items lies in the motives underlying a transaction, which are almost impossible to determine".[2]

SELECTED READINGS

J.R. Hicks, Social Framework, pp. 111-14.

Richard and Giovanna Stone, *National Income and Expenditure*, 9/c, 1971.

Edward Shapiro, *Macroeconomic Analysis*, 4/e, 1978, Ch. 2.

T.F. Dernberg and D.M. McDougall, *Macroeconomics*, 5/e, 1977, Chs. 2 & 3.

John H. Makin, *Macroeconomics*, 1975, Ch. 3.

P. Wonnacott, *Macroeconomics*, 1974, Ch. 17.

W.I. Abraham, *National Income and Economic Accounting*, 1969.

EXERCISES

1. State the essential features of a social accounting model, and point out the main pitfalls which have to be avoided in its use.
2. Give the essentials of Social Accounting and show its use in the study of National Income Flows.
3. What do you understand by social accounting ? How are social accounts arrived ? Discuss the importance of social accounting in

economic analysis.

4. Explain and illustrate social accounting and indicate its usefulness as a tool of economic policy.
5. Explain the input-output transactions analysis of national income acounting.
6. Explain and illustrate flow of funds accounts. In what respects they are different from national income accounts ?
7. Explain fully flow of funds accounts. Discuss their limitations and importance.
8. Show how balance of payments always balances and why does a deficit or surplus arise in the balance of payments of a country ?

---

2. Bo Sodersten and G. Reed, *International Economics*, 3/e, 1994.

CHAPTER 5

THE CIRCULAR FLOW OF INCOME

MEANING

The circular flow of income and expenditure refers to the process whereby the national income and expenditure of an economy flow in a circular manner continuously through time. The various components of national income and expenditure such as saving, investment, taxation, government expenditure, exports, imports, etc. are shown on diagrams in the form of currents and cross-currents in such a manner that national income equals national expenditure.

CIRCULAR FLOW IN A TWO SECTOR ECONOMY

We begin with a simple hypothetical economy where there are only two sectors, the household and business. The household sector owns all the factors of production, that is, land, labour and capital. This sector receives income by selling the services of these factors to the business sector.

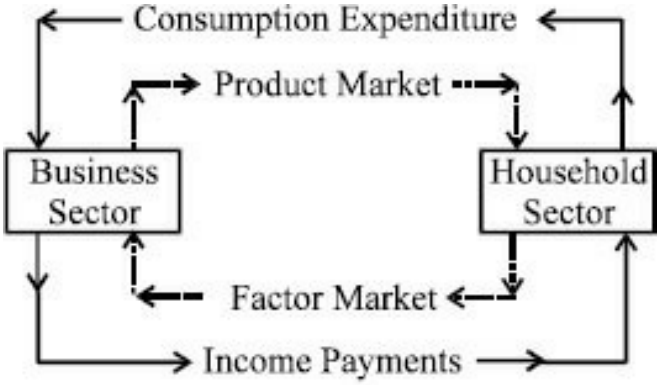


Fig. 1.

The business sector consists of producers who produce products and sell them to the household sector or consumers. Thus the household sector buys the output of products of the business sector. The circular flow of income and expenditure in such an economy is shown in Figure 1 where

the product market is shown in the upper portion and the factor market in the lower portion. In the product market, the household sector purchases goods and services from the business sector while in the factor market the household sector receives income from the former for providing services. Thus the household sector purchases all goods and services provided by the business sector and makes payments to the latter in lieu of these. The business sector, in turn, makes payments to the households for the services rendered by the latter to the business-wage payments for labour services, profit for capital supplied, etc. Thus payments go around in a circular manner from the business sector to the household sector, and from the household sector to the business sector, as shown by arrows in the output portion of the figure. There are also flows of goods and services in the opposite direction to the money payments flows. Goods flow from the business sector to the household sector in the product market, and services flow from the household sector to the business sector in the factor market, as shown in the inner portion of the figure. These two flows give GNP=GNI.

Circular Flow with Saving and Investment Added

The actual economy is not as explained above. In an economy, "inflows" and "leakages" occur in the expenditure and income flows. Such leakages are saving, and inflows or injections are investment which equal each other.

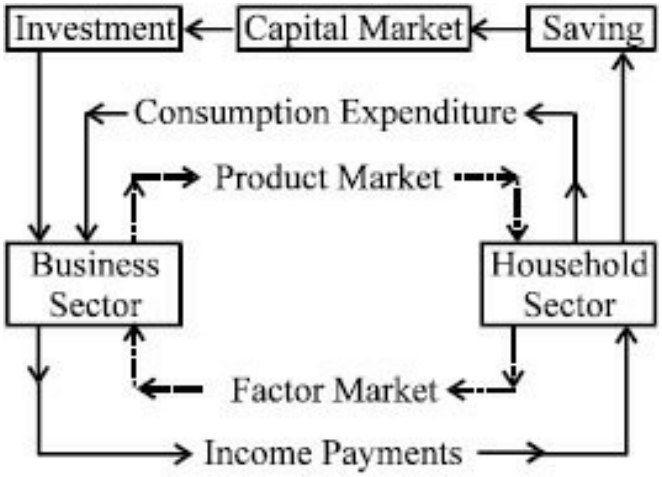


Fig. 2.

Figure 2 shows how the circular flow of income and expenditure is altered by the inclusion of saving and investment. Expenditure has now two alternative paths from household and product markets: (i) directly via consumption expenditure, and (ii) indirectly via investment expenditure.

In Figure 2 there is a capital or credit market in between saving and investment flows from households to business firms. The capital market

refers to a number of financial institutions such as commercial banks, savings banks, loan institutions, the stock and bond markets, etc. The capital market coordinates the saving and investment activities of the households and the business firms. The households supply saving to the capital market and the firms, in turn, obtain investment funds from the capital market.

CIRCULAR FLOW IN A THREE-SECTOR CLOSED ECONOMY

So far we have been working on the circular flow of a two-sector model of an economy. To this we add the government sector so as to make it a three-sector closed model of circular flow of income and expenditure. For this, we add taxation and government purchases(or expenditure) in our presentation. Taxation is a leakage from the circular flow and government purchases are injections into the circular flow.

First, take the circular flow between the household sector and the government sector. Taxes in the form of personal income tax and commodity taxes paid by the household sector are outflows or leakages from the circular flow. But the government purchases the services of the households, makes transfer payments in the form of old age pensions, unemployment relief, sickness benefit, etc., and also spends on them to provide certain social services like education, health, housing, water, parks and other facilities. All such expenditures by the government are injections into the circular flow.

Next take the circular flow between the business sector and the government sector. All types of taxes paid by the business sector to the government are leakages from the circular flow. On the other hand, the government purchases all its requirements of goods of all types from the business sector, gives subsidies and makes transfer payments to firms in order to encourage their production. These government expenditures are injections into the circular flow.

Now we take the household, business and government sectors together to show their inflows and outflows in the circular flow. As already noted, taxation is a leakage from the circular flow.

It tends to reduce consumption and saving of the household sector. Reduced consumption, in turn, reduces the sales and incomes of the firms. On the other hand, taxes on business firms tend to reduce their investment and production. The government offsets these leakages by making purchases from the business sector and buying services of the household sector equal to the amount of taxes. Thus total sales again equal production of firms. In this way, the circular flows of income and expenditure remain in equilibrium.

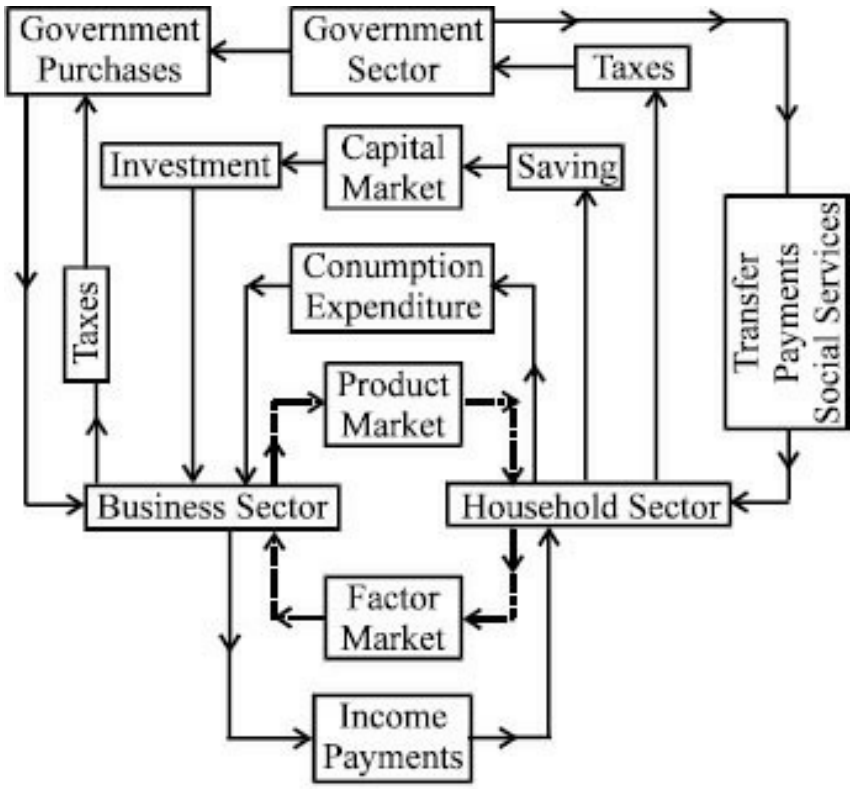


Fig. 3.

Figure 3 shows that taxes flow out of the household and business sectors and go to the government. Now the government makes investment and for this purchases goods from firms and also factors of production from households. Thus government purchases of goods and services are an injection in the circular flow of income and taxes are leakages.

If government purchases exceed net taxes then the government will incur a deficit equal to the difference between the two, i.e., government expenditure and taxes. The government finances its deficit by borrowing from the capital market which receives funds from households in the form of saving. On the other hand, if net taxes exceed government purchases the government will have a budget surplus. In this case, the government reduces the public debt and supplies funds to the capital market which are received by firms.

ADDING FOREIGN SECTOR : CIRCULAR FLOW IN A FOUR-SECTOR OPEN ECONOMY

So far the circular flow of income and expenditure has been shown in the case of a closed economy. But the actual economy is an open one where foreign trade plays an important role. Exports are an injection or inflows into the economy. They create incomes for the domestic firms. When foreigners buy goods and services produced by domestic firms, they are exports in the circular flow of income. On the other hand, imports are leakages from the circular flow. They are expenditures incurred by the household sector to purchase goods from foreign countries. These exports and imports in the circular flow are shown in Figure 4.

Take the inflows and outflows of the household, business and government sectors in relation to the foreign sector. The household sector buys goods imported from abroad and makes payment for them which is a leakage from the circular flow. The households may receive transfer payments from the foreign sector for the services rendered by them in foreign countries.

On the other hand, the business sector exports goods to foreign countries and its receipts are an injection in the circular flow. Similarly, there are many services rendered by business firms to foreign countries such as shipping, insurance, banking, etc. for which they receive payments from abroad. They also receive royalties, interests, dividends, profits, etc. for investments made in foreign countries. On the other hand, the business sector makes payments to the foreign sector for imports of capital goods, machinery, raw materials, consumer goods, and services from abroad. These are the leakages from the circular flow.

Like the business sector, modern governments also export and import goods and services, and lend to and borrow from foreign countries. For all exports of goods, the government receives payments from abroad. Similarly, the government receives payments from foreigners when they visit the country as tourists and for receiving education, etc. and also when the government provides shipping, insurance and banking services to foreigners through the state-owned agencies. It also receives royalities,

interest, dividends etc. for investments made abroad. These are injections into the circular flow. On other hand, the leakages are payments made for the purchase of goods and services to foreigners.

Figure 4 shows the circular flow of the four-sector open economy with saving, taxes and imports shown as leakages from the circular flow on the right hand side of the figure, and investment, government purchases and exports as injections into the circular flow on the left side of the figure. Further, imports, exports and transfer payments have been shown to arise from the three domestic sectors—the household, the business and the government. These outflows and inflows pass through the foreign sector which is also called the "Balance of Payments Sector." If exports exceed imports, the economy has a surplus in the balance of payments. And if imports exceed exports, it has a deficit in the balance of payments. But in the long run, exports of an economy must balance its imports. This is achieved by the foreign trade policies adopted by the economy.

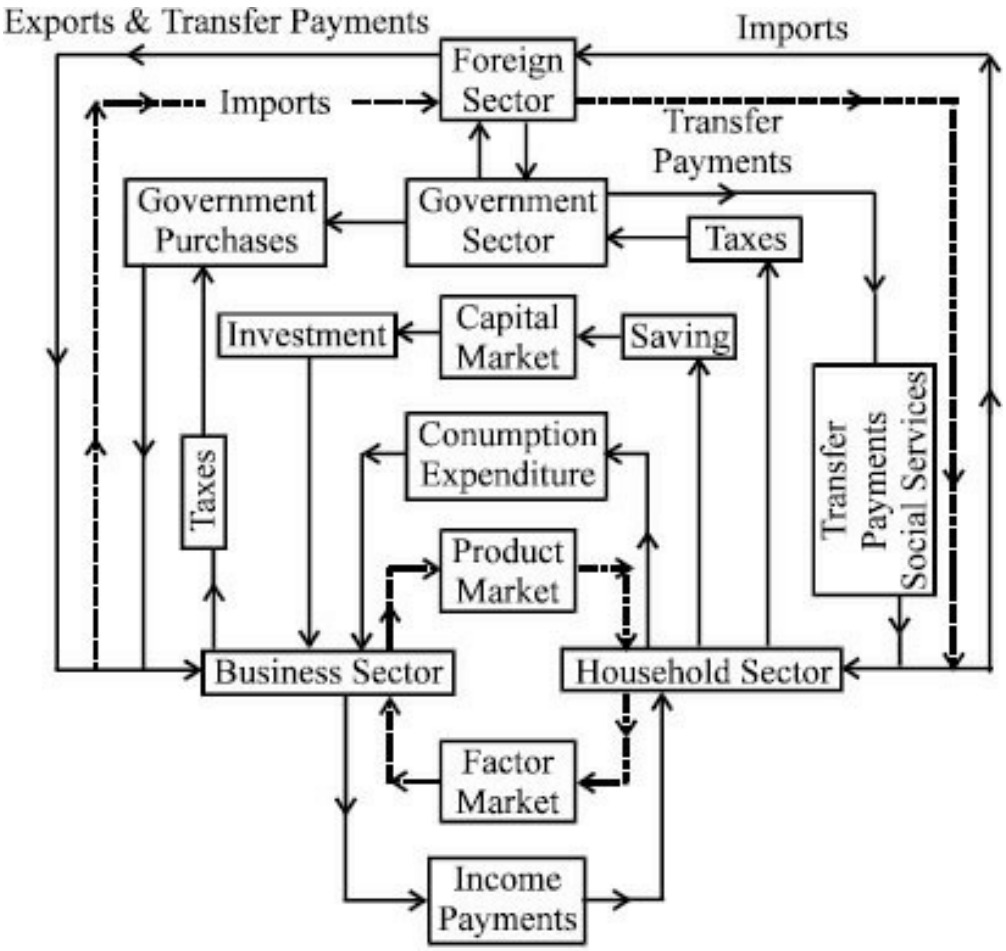


Fig. 4.

The whole analysis can be shown in simple equations :

$$Y = C + I + G \qquad \text{...(1)}$$

where Y represents the production of goods and services, C for consumption expenditure, I investment level in the economy and G for government expenditure respectively.

Now we introduce taxation in the model to equate the government expenditure.

Therefore, $Y = C + S + T$ ...(2)

Where S is saving T is taxation.

By equating (1) and (2), we get

$$C + I + G = C + S + T$$
$$I + G = S + T$$

With the introduction of the foreign sector, we divide investment into domestic investment *(Id)*

and foreign investment (I_f) and get

$$I_d + I_f + G = S + T$$

But $I_f = X - M$

where X is exports and M is imports

$$I_d + (X - M) + G = S + T$$
$$I_d + (X - M) = S + (T - G)$$

The equation shows the equilibrium condition in the circular flow of income and expenditure.

IMPORTANCE OF THE CIRCULAR FLOW

The concept of the circular flow gives a clear-cut picture of the economy. We can know whether the economy is woking efficiently or whether there is any disturbance in its smooth functioning. As such, the circular flow is of immense significance for studying the functioning of the economy and for helping the government in formulating policy measures.

1. Study of Problems of Disequilibrium. It is with the help of circular flow that the problems of disequilibrium and the restoration of equilibrium can be studied.

2. Effects of Leakages and Inflows. The role of leakages enables us to study their effects on the national economy. For example, imports are a leakage out of the circular flow of income because they are payments

made to a foreign country. To stop this leakage, government should adopt appropriate measures so as to increase exports and decrease imports.

3. Link between Producers and Consumers. The circular flow establishes a link between producers and consumers. It is through income that producers buy the services of the factors of production with which the latter, in turn, purchase goods from the producers.

4. Creates a Network of Markets. As a corollary to the above point, the linking of producers and consumers through the circular flow of income and expenditure has created a network of markets for different goods and services where problems relating to their sale and purchase are automatically solved.

5. Inflationary and Deflationary Tendencies. Leakages or injections in the circular flow disturb the smooth functioning of the economy. For example, saving is a leakage out of the expenditure stream. If saving increases, this depresses the circular flow of income. This tends to reduce employment, income and prices, thereby leading to a deflationary process in the economy. On the other hand, consumption tends to increase employment, income, output and prices that lead to inflationary tendencies.

6. Basis of the Multiplier. Again, if leakages exceed injections in the circular flow, the total income becomes less than the total output. This leads to a cumulative decline in employment, income, output, and prices over time. On the other hand, if injections into the circular flow exceed leakages, the income is increased in the economy. This leads to a cumulative rise in employment, income, output, and prices over a period of time. In fact, the basis of the Keynesian multiplier is the cumulative movements in the circular flow of income.

7. Importance of Monetary Policy. The study of circular flow also highlights the importance of monetary policy to bring about the equality of saving and investment in the economy. Figure 2 shows that the equality between saving and investment comes about through the credit or capital market. The credit market itself is controlled by the government through monetary policy. When saving exceeds investment or investment exceeds

saving, money and credit policies help to stimulate or retard investment spending. This is how a fall or rise in prices is also controlled.

8. Importance of Fiscal Policy. The circular flow of income and expenditure points toward the importance of fiscal policy. For national income to be in equilibrium desired saving plus taxes (S+T) must equal desired investment plus government spending (I + G). S+T represent leakages from the spending stream which must be offset by injections of *I + G* into the income stream. If *S + T* exceed *I + G*, government should adopt such fiscal measures as reduction in taxes and spending more itself. On the contrary, If *I + G* exceed *S + T*, the government should adjust its revenue and expenditure by encouraging saving and tax revenue. Thus the circular flow of income and expenditure tells us about the importance of compensatory fiscal policy.

9. Importance of Trade Policies. Similarly, imports are leakages in the circular flow of money because they are payments made to a foreign country. To stop it, the government adopts such measures as to increase exports and decrease imports. Thus the circular flow points toward the importance of adopting export promotion and import control policies.

10. Basis of Flow of Funds Accounts. The circular flow helps in calculating national income on the basis of the flow of funds accounts. The flow of funds accounts are concerned with all transactions in the economy that are accomplished by money transfers. They show the financial transactions among different sectors of the economy, and the link between saving and investment, and lending and borrowing by them.

To conclude, the circular flow of income possesses much theoretical and practical significance in an economy.

EXERCISES

1. Explain the process of circular flow of income and product in a three-sector closed model.
2. Explain the comcept of 'thse circular flow of income'. In what ways do international transactions affect this flow within a closed

economy ?

3. Explain the circular flow of income in an economy including the foreign sector. Use diagrams to illustrate your answer.

PART-III

MACROECONOMIC THEORY

CHAPTER 6

THE CLASSICAL THEORY OF EMPLOYMENT

INTRODUCTION

John Maynard Keynes in his *General Theory of Employment, Interest and Money* published in 1936, made a frontal attack on the classical postulates. He developed a new economics which brought about a revolution in economic thought and policy. The *Gerneral Theory* was written against the background of classical thought. By the "classicists" Keynes meant "the followers of Ricardo, those, that is to say, who adopted and perfected the theory of Ricardian economics." They included, in particular, J.S. Mill, Marshall and Pigou. Keynes repudiated traditional and orthodox economics which had been built up over a century and which dominated economic thought and policy before and during the Great Depression. Since the Keynesian Economics is based on the criticism of classical economics, it is necessary to know the latter as embodied in the theory of employment

THE CLASSICAL THEORY OF EMPLOYMENT

The classical economists believed in the existence of full employment in the economy. To them, full employment was a normal situation and any deviation from this regarded as something abnormal. According to Pigou, the tendency of the economic system is to automatically provide full employment in the labour market when the demand and supply of labour are equal. Unemployment results from the rigidity in the wage structure and interference in the working of free market system in the form of trade union legislation, minimum wage legislation etc. Full employment exists

"when everybody who at the running rate of wages wishes to be employed." Those who are not prepared to work at the existing wage rate are not unemployed because they are voluntarily unemployed. Thus full employment is a situation where there is no possibility of involuntary unemployment in the sense that people are prepared to work at the current wage rate but they do not find work.

The basis of the classical theory is Say's Law of Markets which was carried forward by classical economists like Marshall and Pigou. They explained the determination of output and employment divided into individual markets for labour, goods and money. Each market involves a built-in equilibrium mechanism to ensure full employment in the economy.

Its Assumptions

The classical theory of output and employment is based on the following assumptions :

1. There is the existence of full employment without inflation.

2. There is a *laissez-faire* capitalist economy without government interference.

3. It is a closed economy without foreign trade.

4. There is perfect competition in labour and product markets.

5. Labour is homogeneous.

6. Total output of the economy is divided between consumption and investment expenditures.

7. The quantity of money is given and money is only the medium of exchange.

8. Wages and prices are perfectly flexible.

9. There is perfect information on the part of all market participants.

10. Money wages and real wages are directly related and proportional.

11. Savings are automatically invested and equality between the two is brought about by the rate of interest

12. Capital stock and technical knowledge are given.

13. The law of diminishing returns operates in production.

14. It assumes long run.

Its Explanation

Given these assumptions, the determination of output and employment in the classical theory occurs in labour, goods and money markets in the economy.

Say's Law of Markets

Say's law of markets is the core of the classical theory of employment. An early 19th century French Economist, J.B. Say, enunciated the proposition that "supply creates its own demand." Therefore, there cannot be general overproduction and the problem of unemployment in the economy. If there is general overproduction in the economy, then some labourers may be asked to leave their jobs. The problem of unemployment arises in the economy in the short run. In the long run, the economy will automatically tend toward full employment when the demand and supply of goods become equal. When a producer produces goods and pays wages to workers, the workers, in turn, buy those goods in the market. Thus the very act of supplying (producing) goods implies a demand for them. It is in this way that supply creates its own demand.

Determination of output and Employment

In the classical theory, output and employment are determined by the production function and the demand for labour and the supply of labour in the economy. Given the capital stock, technical knowledge and other factors, a precise relation exists between total output and amount of

employment, i.e., number of workers. This is shown in the form of the following production function:

$$Q = f(K,T,N)$$

where total output *(Q)* is a function *(f)* of capital stock *(K)*, technical knowledge *(T)*, and the number of workers *(N)*.

Given *K* and *T*, the production function becomes $Q = f(N)$ which shows that output is a function of the number of workers. Output is an increasing function of the number of workers, output increases as the employment of labour rises. But after a point when more workers are employed, diminishing marginal returns to labour start. This is shown in Fig. 1 where the curve $Q=f(N)$ is the production function and the total output OQ_1 corresponds to the full employment level N_F. But when more workers $N_F\ N_2$ are employed beyond the full employment level of output OQ_1, the increase in output Q_1Q_2 is less than the increase in employment N_FN_2.

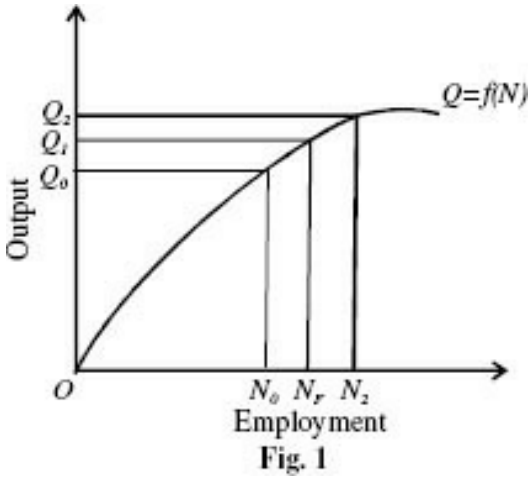


Fig. 1

Labour Market Equilibrium

In the labour market, the demand for labour and the supply of labour determine the level of output and employment. The classical economists regard the demand for labour as the function of the real wage rate: $D_N = f(W/P)$

where D_N = demand for labour, W = wage rate and P = price level. Dividing wage rate *(W)* by price level *(P)*, we get the real wage rate *(W/P)*.

The demand for labour is a decreasing function of the real wage rate, as shown by the downward sloping D_N curve in Fig. 2. It is by reducing the

real wage rate that more workers can be employed.

The supply of labour also depends on the real wage rate : $S_N = f\ (W/P)$, where S_N is the supply of labour. But it is an increasing function of the real wage rate, as shown by the upward sloping S_N curve in Fig. 2. It is by increasing the real wage rate that more workers can be employed.

When the D_N and S_N curves intersect at point *E*, the full employment level N_F is determined at the equilibrium real wage rate W/P_0. If the wage rate rises from WP_0 to WP_1, the supply of labour will be more than its demand by *ds*. Now at W/P_1 wage rate, ds workers will be involuntary unemployed because the demand for labour (W/P_1-*d*) is less than their supply (W/P_1-*s*). With competition among workers for work, they will be willing to accept a lower wage rate. Consequently, the wage rate will fall from W/P_1 to W/P_0 . The supply of labour will fall and the demand for labour will rise and the equilibrium point *E* will be restored along with the full employment level N_F. On the contrary, if the wage rate falls from W/P_0 *to* WP_2 the demand for labour (W/P_2-d_1) will be more than its supply (W/P_2-s_1). Competition by employers for workers will raise the wage rate from W/P_2 to W/P_0 and the equilibrium point *E* will be restored along with the full employment level N_F.

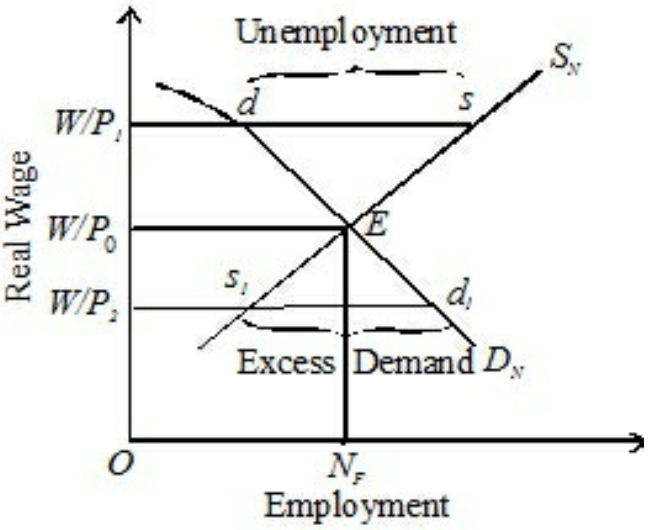


Fig. 2

Wage Price Flexibility

The classical economists believed that there was always full employment in the economy. In case of unemployment, a general cut in money wages would take the economy to the full employment level. This argument is

based on the assumption that there is a *direct* and *proportional* relation between money wages and real wages. When money wages are reduced, they lead to reduction in cost of production and consequently to the lower prices of products. When prices fall, demand for products will increase and sales will be pushed up. Increased sales will necessitate the employment of more labour and ultimately full employment will be attained.

Pigou explains the entire proposition in the equation : $N = qY/W$. In this equation, N is the number of workers employed, q is the fraction of income earned as wages, Y is the national income and W is the money wage rate. N can be increased by a reduction in W. Thus the key to full employment is a reduction in money wage. When prices fall with the reduction of money wage, real wage is also reduced in the same proportion.

As explained above, the demand for labour is a decreasing function of the real wage rate. If W is the money wage rate, P is the price of the product, and MP_N is the marginal product of labour, we have

$$W = P \sqcap MP_N \quad \text{or} \quad W/P = MP_N$$

Since MP_N declines as employment increases, it follows that the level of employment increases as the real wage (W/P) declines. This is explained in Figure 3. In Panel (A), S_N is the supply curve of labour and D_N is the demand curve for labour. The intersection of the two curves at E shows the level of full employment N_F and the real wage W/P_0. If the real wage rises to W/P_1, supply exceeds the demand for labour by sd and N_1N_2 workers are unemployed. It is only when the wage is reduced to W/P_0 that unemployment disappears and the level of full employment is attained.This is shown in Panel (B), where MP_N is the marginal product of labour curve which slopes downward as more labour is employed. Since every worker is paid wages equal to his marginal product, therefore the full employment level N_F is reached when the wage rate falls from W/P_1 to W/P_0. Contrariwise, with the fall in the wage from W/P_0 to W/P_2, the demand for labour increases more than its supply by s_1d_1, the workers

demand higher wage. This leads to the rise in the wage from W/P_2 to W/P_0 and the full employment level N_F is attained.

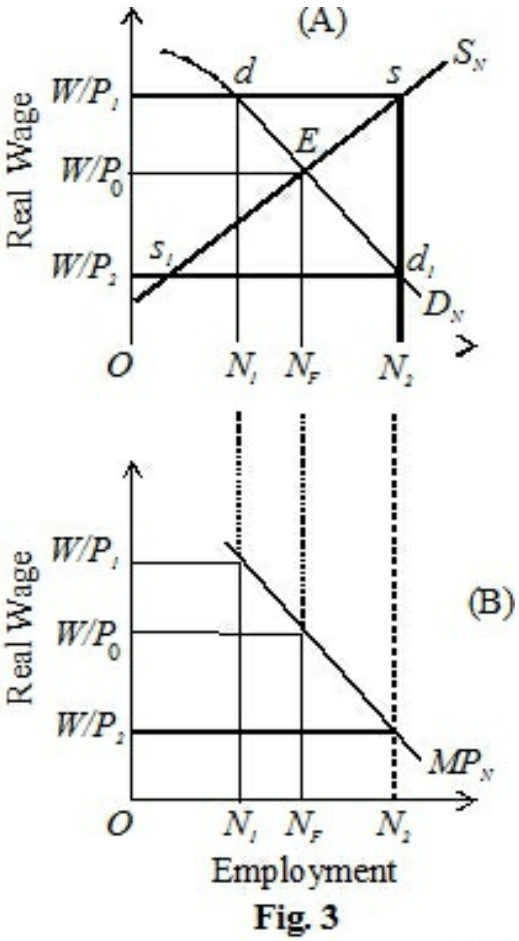


Fig. 3

Goods Market Equilibrium

The goods market is in equilibrium when saving equals investment. At that point of time, total demand equals total supply and the economy is in a state of full employment. According to the classicists, what is not spent is automatically invested. Thus saving must equal investment. If there is any divergence between the two, the equality is maintained through the mechanism of the rate of interest. To them, both saving and investment are the functions of the interest rate,

$$S = f(r) \quad \text{...(1)}$$
$$I = f(r) \quad \text{...(2)}$$
$$S = I$$

where S = saving, I = investment, and r = interest rate.

To the classicists, interest is a reward for saving. The higher the rate of interest, the higher the saving, and lower the investment. On the contrary, the lower the rate of interest, the higher the demand for investment funds, and lower the saving. If at any given period, investment exceeds saving, ($I>S$) the rate of interest will rise. Saving will increase and investment will decline till the two are equal at the full employment level. This is because saving is regarded as an increasing function of the interest rate and investment as a decreasing function of the rate of

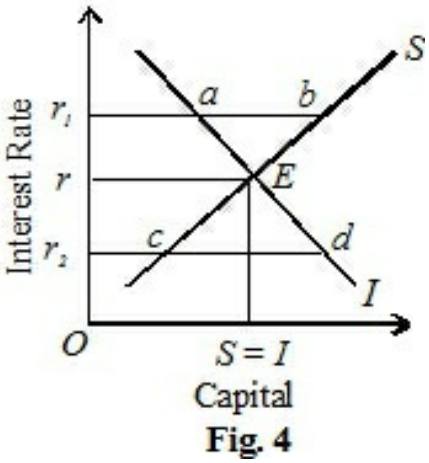


Fig. 4

interest.

Assuming interest rates are perfectly elastic, the mechanism of the equality between saving and investment is shown in Figure 4 where *S* is the saving curve and *I* is the investment curve. Both intersect at *E* which is the full employment level where at *Or* interest rate *S* = I. If the interest rate rises to Or_1, saving is more than investment by *ba* which will lead to unemployment in the economy. Since *S* > I, the investment demand for capital being less than its supply, the interest rate will fall to Or, investment will increase and saving will decline. Consequently, *S* = *I* equilibrium will be re-established at point *E.* On the contrary, with a fall in the interest rate from *Or* to Or_2, investment will be more than saving (I > S) by cd, the demand for capital will be more than its supply. The interest rate will rise, saving will increase and investment will decline. Ultimately, *S* = *I* equilibrium will be restored at the full employment level E.

Money Market Equilibrium

The money market equilibrium in the classical theory is based on the Quantity Theory of Money which states that the general price level *(P)* in the economy depends on the supply of money (M). The equation is *MV* = PT, where *M* = supply of money, *V* = velocity of circulation of M, *P* = Price level, and *T* = volume of transaction or total output.

The equation tells that the total money supply *MV* equals the total value of output *PT* in the economy. Assuming *V* and *T* to be constant, a change in the supply of money *(M)* causes a proportional change in the price level *(P).* Thus the price level is a function of the money supply : *P = f (M)*.

The relation between quantity of money, total output and price level is depicted in Figure 5 where the price level is taken on the horizontal axis and the total output on the vertical axis. *MV* is the money supply curve which is a rectangular hyperbola. This is because the equation *MV = PT* holds on all points of this curve. Given the output level *OQ,* there would be only one price level *OP* consistent with the quantity of money, as shown by point *M* on the *MV* curve. If the quantity of money increases,

the *MV* curve will shift to the right as M_1V curve. As a result, the price level would rise from *OP* to OP_1, given the same level of output *OQ*.

This rise in the price level is exactly proportional to the rise in the quantity of money, i.e., $PP_1 = MM_1$ when the full employment level of output remains *OQ*.

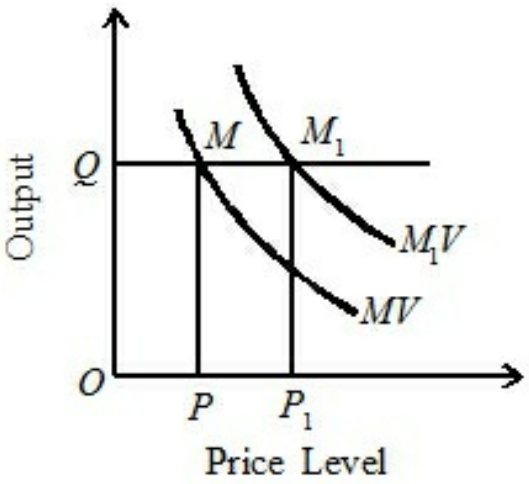


Fig. 5.

COMPLETE CLASSICAL MODEL – A SUMMARY

The classical theory of employment was based on the assumption of full employment where full employment was a normal situation and any deviation from this was regarded as an abnormal situation. This was based on Say's Law of Market. According to this, supply creates its own demand and the problem of overproduction and unemployment does not arise. Thus there is always full employment in the economy. If there is overproduction and unemployment, the automatic forces of demand and supply in the market will bring back the full employment level.

In the classical theory, the determination of output and employment takes place in labour, goods and money markets of the economy, as shown in Fig. 6. The forces of demand and supply in these markets will ultimately bring full employment in the economy.

In the classical analysis, output and employment in the economy are determined by the aggregate production function, demand for labour and supply of labour. Given the stock of capital, technical knowledge and other factors, there is a precise relation between total output and employment (number of workers). This is expressed as $Q = f(K,T,N)$. In other words, total output *(Q)* is a function *(f)* of capital stock (K), technical knowledge *T,* and number of workers *(N).* Given *K* and T, total output *(Q)* is an increasing function of the number of workers *(N)* : $Q = f$

(N) as shown in Panel (B). At point E, ON_F workers produce *OQ* output. But beyond point E, as more workers are employed, diminishing marginal returns start.

Labour Market Equilibrium. In the labour market, the demand for and supply of labour determine output and employment in the economy. The demand for labour depends on total output. As production increases, the demand for labour also increases. The demand for labour, in turn, depends on the marginal productivity *(MP)* of labour which declines as more workers are employed. The supply of labour depends on the wage rate, $S_L = f$ (W/P), and is an increasing function of the wage rate. The demand for labour also depends on the wage rate, $D_L = f\ (W/P)$, and is a decreasing function of the wage rate. Thus both the demand for and supply of labour are the functions of real wage rate *(W/P).* The intersection point *E* of D_L and S_L curves at *W/P* wage rate in Panel (C) of the figure determines the full employment level ON_F.

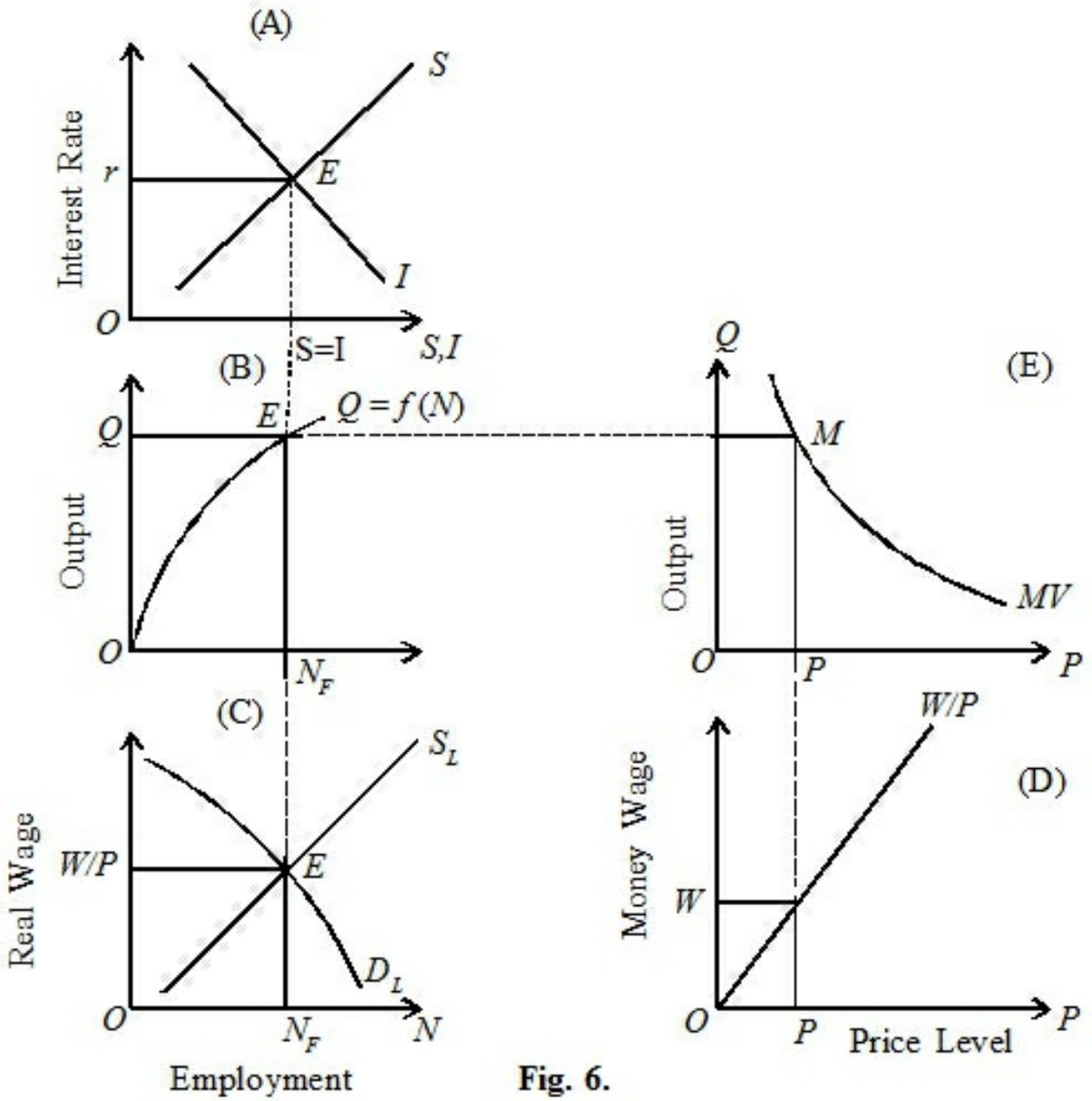


Fig. 6.

Goods Market Equilibrium. In the classical analysis, the goods market is in equilibrium when saving and investment are in equilibrium *(S=I).* This equality is brought about by the mechanism of interest rate at the full employment level of output so that the quantity of goods demanded is equal to the quantity of goods supplied. This is shown in Panel (A) of the figure where S=I at point E when the interest rate is *Or.*

Money Market Equilibrium. The money market is in equilibrium when the demand for money equals the supply of money. This is explained by

the Quantity Theory of Money which states that the quantity of money is a function of the price level, $P = f(MV)$. Changes in the general price level are proportional to the quantity of money. The equilibrium in the money market is shown by the equation $MV = PT$ where MV is the supply of money and PT is the demand for money. The equilibrium of the money market explains the price level corresponding to the full employment level of output which relates Panel (E) and Panel (B) with MQ line.

The price level OP is determined by total output (Q) and the quantity of money (MV), as shown in Panel (E). Then the real wage corresponding with the money wage is determined by the (W/P) curve, as shown in Panel (D). When the money wage increases, the real wage also increases in the same proportion and there is no effect on the level of output and employment. It follows that the money wage should be reduced in order to attain the full employment level in the economy. Thus the classicists favoured a flexible price-wage policy to maintain full employment.

KEYNES'S CRITICISM OF CLASSICAL THEORY

Keynes vehemently criticised the classical theory of employment for its unrealistic assumptions in his *General Theory*. He attacked the classical theory on the following counts:

(1) Underemployment Equilibrium. Keynes rejected the fundamental classical assumption of full employment equilibrium in the economy. He considered it as unrealistic. He regarded full employment as a special situation. The general situation in a capitalist economy is one of underemployment. This is because the capitalist society does not function according to Say's law, and supply always exceeds its demand. We find millions of workers are prepared to work at the current wage rate, and even below it, but they do not find work. Thus the existence of involuntary unemployment in capitalist economies (entirely ruled out by the classicists) proves that underemployment equilibrium is a normal situation and full employment equilibrium is abnormal and accidental.

(2) Refutation of Say's Law. Keynes refuted Say's Law of markets that supply always created its own demand. He maintained that all income

earned by the factor owners would not be spent in buying products which they helped to produce. A part of the earned income is saved and is not automatically invested because saving and investment are distinct functions. So when all earned income is not spent on consumption goods and a portion of it is saved, there results in a deficiency of aggregate demand. This leads to general overproduction because all that is produced is not sold. This, in turn, leads to general unemployment. Thus Keynes rejected Say's Law that supply created its own demand. Instead he argued that it was demand that created supply. When aggregate demand rises, to meet that demand, firms produce more and employ more people.

(3) Self-adjustment not Possible. Keynes did not agree with the classical view that the *laissez-faire* policy was essential for an automatic and self-adjusting process of full employment equilibrium. He pointed out that the capitalist system was not automatic and self-adjusting because of the non-egalitarian structure of its society. There are two principal classes, the rich and the poor. The rich possess much wealth but they do not spend the whole of it on consumption. The poor lack money to purchase consumption goods. Thus there is general deficiency of aggregate demand in relation to aggregate supply which leads to overproduction and unemployment in the economy. This, in fact, led to the Great Depresion. Had the capitalist system been automatic and self-adjusting, this would not have occurred. Keynes, therefore, advocated state intervention for adjusting supply and demand within the economy through fiscal and monetary measures.

(4) Equality of Saving and Investment through Income Changes. The classicists believed that saving and investment were equal at the full employment level and in case of any divergence, the equality was brought about by the machanism of rate of interest. Keynes held that the level of saving depended upon the level of income and not on the rate of interest. Similarly investment is determined not only by rate of interest but by the marginal efficiency of capital. A low rate of interest cannot increase investment if business expectations are low. If saving exceeds investment, it means people are spending less on consumption. As a result, demand declines. There is overproduction and fall in investment, income, employment and output. It will lead to reduction in saving and ultimately

the equality between saving and investment will be attained at a lower level of income. Thus it is variations in income rather than in interest rate that bring the equality between saving and investment.

(5) Importance of Speculative Demand for Money. The classical economists believed that money was demanded for transactions and precautionary purposes. They did not recognise the speculative demand for money because money held for speculative purposes related to idle balances. But Keynes did not agree with this view. He emphasised the importance of speculative demand for money. He pointed out that the earning of interest from assets meant for transactions and precautionary purposes may be very small at a low rate of interest. But the speculative demand for money would be infinitely large at a low rate of interest. Thus the rate of interest will not fall below a certain minimum level, and the speculative demand for money would become perfectly interest elastic. This is Keynes 'liquidity trap' which the classicists failed to analyse.

(6) Rejection of Quantity Theory of Money. Keynes rejected the classical Quantity Theory of Money on the ground that increase in money supply will not necessarily lead to rise in prices. It is not essential that people may spend all extra money. They may desposit it in the bank or save. So the velocity of circulation of money (V) may slow down and not remain constant. Thus V in the equation MV = PT may vary. Moreover, an increase in money supply, may lead to increase in investment, employment and output if there are idle resources in the economy and the price level (P) may not be affected.

(7) Money not Neutral. The classical economists regarded money as neutral. Therefore, they excluded the theory of output, employment and interest rate from monetary theory. According to them, the level of output and employment and the equilibrium rate of interest were determined by real forces. Keynes criticised the classical view that monetary theory was separate from value theory. He integrated monetary theory with value theory, and brought the theory of interest in the domain of monetary theory by regarding the interest rate as a monetary phenomenon. He integrated the value theory and the monetary theory through the theory of output. This he did by forging a link between the quantity of money and

the price level via the rate of interest. For instance, when the quantity of money increases, the rate of interest falls, investment increases, income and output increase, demand increases, factor costs and wages increase, relative prices increase, and ultimately the general price level rises. Thus Keynes integrated monetary and real sectors of the economy.

(8) Refutation of Wage-Cut. Keynes refuted the Pigovian formulation that a cut in money wage could achieve full employment in the economy. The greatest fallacy in Pigou's analysis was that he extended the argument to the economy which was applicable to a particular industry. Reduction in wage rate can increase employment in an industry by reducing costs and increasing demand. But the adoption of such a policy for the economy leads to a reduction in employment. When there is a general wage-cut, the income of the workers is reduced. As a result, aggregate demand falls leading to a decline in employment.

From the practical view point also Keynes never favoured a wage cut policy. In modern times, workers have formed strong trade unions which resist a cut in money wage. They would resort to strikes. The consequent unrest in the economy would bring a decline in output and income. Moreover, social justice demands that wages should not be cut if profits are left untouched.

(9) No Direct and Proportionate Relation between Money and Real Wages. Keynes also did not accept the classical view that there was a direct and proportionate relationship between money wages and real wages. According to him, there is an inverse relation between the two. When money wages fall, real wages rise and vice versa. Therefore, a reduction in the money wage would not reduce the real wage, as the classicists believed, rather it would increase it. This is because the money wage cut will reduce cost of production and prices by more than the former. Thus the classical view that fall in real wages will increase employment breaks down. Keynes, however, believed that employment could be increased more easily through monetary and fiscal measures rather than by reduction in money wage. Moreover, institutional resistances to wage and price reductions are so strong that it is not possible to implement such a policy administratively.

(10) State Intervention Essential. Keynes did not agree with Pigou that "frictional maladjustments alone account for failure to utilise fully our productive power." The capitalist system is such that left to itself it is incapable of using productive power fully. Therefore, state intervention is necessary. The state may directly invest to raise the level of economic activity or to supplement private investment. It may pass legislation recognising trade unions, fixing minimum wages and providing relief to workers through social security measures. "Therefore", as observed by Dillard, "it is bad politics even if it should be considered good economics to object to labour unions and to liberal labour legislation." So Keynes favoured state action to utilise fully the resources of the economy for attaining full employment.

(11) Long-Run Analysis Unrealistic. The classicists believed in the long-run full employment equilibrium through a self-adjusting process. Keynes had no patience to wait for the long period for he believed that "In the long-run we are all dead". As pointed by Schumpeter, "His philosophy of life was essentially a short-term philosophy." His analysis is confined to short-run phenomena. Unlike the classicists, he assumes tastes, habits, techniques of production, supply of labour, etc. to be constant during the short period and so neglects long-run influences on demand. Assuming consumption demand to be constant, he lays emphasis on increasing investment to remove unemployment. But the equilibrium level so reached is one of underemployment rather than of full employment.

Thus the classical theory of employment is unrealistic and is incapable of solving the present day economic problems of the capitalist world.

SELECTED READING

D. C. Rowan, *Output, Inflation and Growth,* 2/e, 1974, Ch. 18.

Michael C. Lovell, *Macroeconomics,* 1975. Ch. 12.

Paul Wonnacott, *Macroeconomics,* 1974. Ch. 2.

Edward Shapiro, *Macroeconomics Analysis,* 4/e, 1978, Ch. 14.

E. E. Hagen, "The Classical Theory of the Level of Output and Employment", in M.G. Mueller (ed.), *Readings in Macroeconomics*, 2/e, 1971, kCh. 1.

W.H. Branson, *Macroeconomic Theory and Policy*, 1972, Ch. 6.

EXERCISES

1. Why did the classical economists assume that in a competitive economy full employment will be automatic? On what grounds Keynes questioned this assumption?
2. Explain the classical theory of employment. On what grounds Keynes criticised it ?

CHAPTER 7

SAY'S LAW OF MARKET

SAY'S LAW

Say's law of markets is the core of the classical theory of employment. An early 19th century French Economist, J.B. Say, enunciated the proposition that "supply creates its own demand." Therefore, there cannot be general overproduction and the problem of unemployment in the economy. On the other hand, if there is general overproduction in the economy, then some labourers may be asked to leave their jobs. There may be the problem of unemployment in the economy for sometime. In the long-run, the economy will automatically tend toward full employment. In Say's words, "It is production which creates markets for goods. A product is no sooner created than it, from that instant, affords a market for other products to the full extent of its own value. Nothing is more favourable to the demand of one product, than the supply of another."[1] This definition explains the following important facts about the law.

Production Creates Market (Demand) for Goods. When producers obtain the various inputs to be used in the production process, they generate the necessay income. For example, producers give wages to labourers for producing goods. The labourers will purchase the goods from the market for their own use. This, in turn, causes the demand for goods produced. In this way, supply creates its own demand.

---

1. *J.B. Say,* Political Economy, *4/e, Vol.7 pp.167-170*

Barter System as its Basis. In its original form, the law is applicable to a

barter economy where goods are ultimately sold for goods. Therefore, whatever is produced is ultimately consumed in the economy. In other words, people produce goods for their own use to sustain their consumption levels. Say's law, in a very broad way, is, as Prof. Hansen has said, "a description of a free-exchange economy. So conceived, it illuminates the truth that the main source of demand is the flow of factor income generated from the process of production itself.[2] Thus, the existence of money does not alter the basic law.

General Overproduction Impossible. If the production process is continued under normal conditions, then there will be no difficulty for the producers to sell their products in the market. According to Say, work being unpleasant, no person will work to make a product unless he wants to exchange it for some other product which he desires. Therefore, the very act of supplying goods implies a demand for them. In such a situation, there cannot be general overproduction because supply of goods will not exceed demand as a whole. But a particular good may be over produced because the producer incorrectly estimates the quantity of the product which others want. But this is a temporary phenomenon, for the excess production of a particular product can be corrected in time by reducing its production.

J.S. Mill supported Say's views regarding the impossibility of general overproduction and general unemployment. According to him, Say's law of markets does not consider the possibility of general overproduction and also rejects the possibility of decrease in the demand of goods produced in the economy. By employing more factors of production, there is an increase in the level of employment and therefore profits are maximised.

Saving-Investment Equality. Income accruing to the factor owners in the form of rent, wages and interest is not spent on consumption but some proportion out of it is saved which is automatically invested for further production. Therefore, investment in production is a saving which helps to create demand for goods in the market. Further, saving-investment equality is maintained to avoid general overproduction.

Rate of Interest as a Determinant Factor. Say's law of markets regards the rate of interest as a determinant factor in maintaining the equality

between saving and investment. If there is any divergence between the two, the equality is maintained through the mechanism of the rate of interest. If at any given time investment exceeds saving, the rate of interest will rise. To maintain the equality, saving will increase and investment will decline. This is due to the fact that saving is regarded as an increasing function of the interest rate, and investment as a decreasing function of the rate of interest. On the contrary, when saving is more than investment, the rate of interest falls, investment increases and saving declines till the two are equal at the new interest rate.[3]

Labour Market. Prof. Pigou formulated Say's law in terms of labour market. By giving minimum wages to labourers, according to Pigou, more labourers can be employed. In this way, there will be more demand for labour. As pointed out by Pigou, "with perfectly free competition...there will always be at work a strong tendency for wage rates to be so related to demand that everybody is employed."[4] Unemployment results from rigidity in the wage structure and interferences in the working of the free market economy. Direct interference comes in the form of minimum wage laws passed by the state. The trade unions may be demanding higher wages, more facilities and reduction in working hours. In short, it is only under free competition that the tendency of the economic system is to provide automatically full employment in the labour market.[5]

---

2. *A.H. Hansen,* A Guide to Keynes, *p.3*

3. *Give Fig. 4 of Ch. 6*

PROPOSITIONS AND IMPLICATIONS OF THE LAW

Say's propositions and its implications present the true picture of the market law. These are given below.

1. Full Employment in the Economy. The law is based on the proposition that there is full employment in the economy. Increase in production means more employment to the factors of production. Production continues to increase until the level of full employment is reached. Under such a situation, the level of production will be maximum.

2. Proper Utilization of Resources. If there is full employment in the economy, idle resources will be properly utilized which will further help to produce more and also generate more income.

3. Perfect Competition. Say's law of market is based on the proposition of perfect competition in labour and product markets. Other conditions of perfect competition are given below:

(a) Size of the Market. According to Say's law, the size of the market is large enough to create demand for goods. Moreover, the size of the market is also influenced by the forces of demand and supply of various inputs.

(b) Automatic Adjustment Mechanism. The law is based on this proposition that there is automatic and self-adjusting mechanism in different markets. Disequilibrium in any market is a temporary situation. For example, in capital market, the equality between saving and investment is maintained by the rate of interest while in the labour market the adjustment between demand and supply of labour is maintained by the wage rate.

(c) Role of Money as Neutral. The law is based on the proposition of a barter system where goods are exchanged for goods. But it is also assumed that the role of money is neutral. Money does not affect the production process.

4. Laissez-faire Policy. The law assumes a closed capitalist economy which follows the policy of *laissez-faire*. The policy of laissez-faire is essential for an automatic and self-adjusting process of full employment equilibrium.

5. Saving as a Social Virtue. All factor income is spent in buying goods which they help to produce. Whatever is saved is automatically invested for further production. In other words, saving is a social virtue.

---

4. *A.C. Pigou*, Theory of Employment., *p.252*

5. *Give Fig. 3 of Ch. 6*

CRITICISMS OF SAY'S LAW

J.M. Keynes in his *General Theory* made a frontal attack on the classical postulates and Say's law of markets. He criticised Say's law of markets on the following grounds:

1. Supply does not Create its Demand. Say's law assumes that production creates market(demand) for goods. Therefore, supply creates its own demand. But this proposition is not applicable to modern economies where demand does not increase as much as production increases. It is also not possible to consume only those goods which are produced within the economy.

2. Self-adjustment not Possible. According to Say's law, full-employment is maintained by an automatic and self-adjustment mechanism in the long run. But Keynes had no patience to wait for the long period for he believed that "In the long-run we are all dead." It is not the automatic adjustment process which removes unemployment. But unemployment can be removed by increase in the rate of investment.

3. Money is not Neutral. Say's law of markets is based on a barter system and ignores the role of money in the system. Say believes that money does not affect the economic activities of the markets. On the other hand, Keynes has given due importance to money. He regards money as a medium of exchange. Money is held for income and business motives. Individuals hold money for unforeseen contingencies while businessmen keep cash in reserve for future activities.

4. Over Production is Possible. Say's law is based on the proposition that supply creates its own demand and there cannot be general over-production. But Keynes does not agree with this proposition. According to him, all income accruing to factors of production is not spent but some fraction out of it is saved which is not automatically invested. Therefore, saving and investment are always not equal and it becomes the problem of overproduction and unemployment.

5. Underemployment Situation. Keynes regards full employment as a

special case because there is underemployment in capitalist economies. This is because the capitalist economies do not function according to Say's law and supply always exceeds its demand. For example, millions of workers are prepared to work at the current wage rate, and even below it, but they do not find work.

6. State Intervention. Say's law is based on the existence of laissez-faire policy. But Keynes has highlighted the need for state intervention in the case of general overproduction and mass unemployment. *Laissez-faire*, infact, led to the Great Depression. Had the capitalist system been automatic and self-adjusting. This would not have occurred. Keynes, therefore, advocated state intervention for adjusting supply and demand within the economy through fiscal and monetary measures.

7. Equality through Income. Keynes does not agree with the classical view that the equality between saving and investment is brought about through the mechanism of interest rate. But in reality, it is changes in income rather than the rate of interest which bring the two to equality.

8. Wage-cut no Solution. Pigou favoured the policy of wage-cut to solve the problem of unemployment. But Keynes opposed such a policy both from the theoretical and practical points of view. Theoretically, a wage-cut policy increases unemployment instead of removing it. Practically, workers are not prepared to accept a cut in money wage. Keynes, therefore, favoured a flexible monetary policy to a flexible wage policy to raise the level of employment in the economy.

9. Demand Creates its own supply. Say's law of market is based on the proposition that "supply creates its own demand". Therefore, there cannot be general overproduction and mass unemployment. Keynes has criticised this proposition and propounded the opposite view that *demand creates its own supply*. Unemployment results from the deficiency of effective demand because people do not spend the whole of their income on consumption.

EXERCISES

1. "Supply creates its own demand". Critically examine this statement.
2. "General overproduction and general unemployment are impossible". Discuss.
3. State and explain the main propositions of Say's law. On what grounds Keynes criticised them ?

CHAPTER 8

THE PRINCIPLE OF EFFECTIVE DEMAND : AGGREGATE DEMAND AND AGGREGATE SUPPLY

INTRODUCTION

The logical starting point of Keynes's theory of employment is the principle of effective demand. In a capitalist economy, the level of employment depends on effective demand. Thus unemployment results from a deficiency of effective demand and the level of employment can be raised by increasing the level of effective demand.

EFFECTIVE DEMAND

In ordinary parlance, demand means desire. It becomes effective when income is spent in buying consumption goods and investment goods. Keynes used the term 'effective demand' to denote the total demand for goods and services at various levels of employment. Different levels of employment represent different levels of aggregate demand. But there can be a level of employment where aggregate demand equals aggregate supply. This is the point of effective demand. In Keynes's words, "The value of *D* (Aggregate Demand) at the point of Aggregate Demand function, where it is intersected by the Aggregate Supply function, will be called the effective demand." Thus according to Keynes, the level of employment is determined by effective demand which, in turn, is determined by aggregate demand price and aggregate supply price.

Aggregate Demand Price

"The aggregate demand price for the output of any given amount of employment is the total sum of money or proceeds, which is expected from the sale of the output produced when that amount of labour is employed."[1] Thus the aggregate demand price is the amount of money which the entrepreneurs *expect* to get by selling the ouput produced by the number of men employed. In other words, it refers to the *expected revenue* from the sale of output produced at a particular level of employment. Different aggregate demand prices relate to different levels of employment in the economy.

A statement showing the various aggregate demand prices at different levels of employment is called the aggregate demand price schedule or aggregate demand function. "The aggregate demand function," according to Keynes, "relates any given level of employment to the expected proceeds from that level of employment." Table I shows the aggregate demand schedule.

The table reveals that with the increase in the level of employment proceeds expected rise and at lower levels of employment decline. When 45 lakh people are provided employment the aggregate demand price is Rs 280 crores and when 25 lakh people are provided jobs, it is Rs 240 crores. According to Keynes, the aggregate demand function is an increasing function of the level of employment and is expressed as $D = F(N)$, where D is the proceeds which entrepreneurs expect from the employment of N men.

Table I: Aggregate Demand Schedule

Level of Employment(N) (in Lakhs)	*Aggregate Demand Price(D)* (Rs Crores)
20	230
25	240
30	250
35	260

40	270
45	280
50	290

The aggregate demand curve can be drawn on the basis of the above schedule. It slopes upward from left to right because as the level of employment increases aggregate demand price also rises, shown as *AD* curve in Figure 1.

Aggregate Supply Price

When an entrepreneur gives employment to certain amount of labour, it requires certain quantities of cooperant factors like land, capital, raw materials, etc. which will be paid remuneration along with labour. Thus each level of employment involves certain money costs of production including normal profits which the entrepreneur must cover. "At any given level of employment of labour aggregate supply price is the total amount of money which all the entrepreneurs in the economy, taken together, must expect to receive from the sale of the output produced by that given number of men, if it is to be just worth employing them."[2] In brief, the aggregate supply price refers to the *proceeds necessary* from the sale of output at a particular level of employment. Thus each level of employment in the economy is related to a particular aggregate supply price and there are different aggregate supply prices for different levels of employment.

---

1. *D.Dillard*, op. cit.

A statement showing the various aggregate supply prices at different levels of employment is called the aggregate supply price schedule or aggregate supply function. In the words of Prof. Dillard, "The aggregate supply function is a schedule of the minimum amounts of proceeds required to induce varying quantities of employment."[3] Table II shows the aggregate supply schedule.

Table II: Aggregate Supply Schedule

Level of Employment(N) (in lakhs)	*Aggregate Supply Price(Z) (Rs crores)*
20	215
25	230
30	245
35	260
40	275
40	290
40	305

The above table reveals that the aggregate supply price rises with the increase in the level of employment. If entrepreneurs are to provide employment to 20 lakh workers, they must receive Rs 215 crores from the sale of the output produced by them. It is only when they expect to receive the minimum amounts of proceeds(Rs 230 crores, Rs 245 crores and Rs 260 crores) that they will provide employment to more workers(25 lakhs, 30 lakhs and 35 lakhs respectively). But when the economy reaches the level of full employment(at 40 lakh workers) the aggregate supply price(Rs 275,290 and 305 crores) continues to increase but there is no further increase in employment. According to Keynes, the aggregate supply function is an increasing function of the level of employment and is expressed as $Z = \varphi N$, where Z is aggregate supply price of the output from employing N men.

The aggregate supply curve can be drawn on the basis of the schedule. It slopes upward from left to right because as the necessay expected proceeds increase, the level of employment also rises. But when the economy reaches the level of full employment, the aggregate supply curve becomes vertical. Even with the increase in the aggregate supply price, it is not possible to provide more employment as the economy has attained the level of full employment.

---

2. *Stonier and Hague,* A Text Book of Economic Theory, *p. 381*

3. *D.Dillard,* op. cit., *p.31.*

Determination of Effective Demand

We have studied the two determinants of effective demand separately and now are in a position to analyse the process of determining the level of employment in the economy. The level of employment is determined at the point where the aggregate demand price equals the aggregate supply price. In other words, it is the point where what entrepreneurs *expect to receive* equals what they *must receive* and their profits are maximised. This point is called the *effective demand* and here entrepreneurs earn normal profits. So long as the aggregate demand price is higher than the aggregate supply price, the prospects of getting additional profits are greater when more workers are provided employment. The proceeds expected(revenue) rise more than the proceeds necessary(costs). This process will continue till the aggregate demand price equals the aggregate supply price and the point of effective demand is reached. This point determines the level of employment and output in the economy. The point of effective demand is, however, not necessarily one of full employment but of underemployment equilibrium. If the entrepreneurs try to provide more employment after this point, the aggregate supply price exceeds the aggregate demand price indicating that the total costs are higher than the total revenue and there are losses. So the entrepreneurs will not employ workers beyond the point of effective demand till the aggregate demand price rises to meet the aggregate supply price at the new equilibrium point which may be one of full employment. If the aggregate demand price is raised still further, it will lead to inflation because no increase in employment and output is possible beyond the level of full employment. The following table explains the determination of the point of effective demand.

Table III shows that so long as the aggregate demand price is higher than the aggregate supply price, it is profitable for entrepreneurs to employ more workers, when they expect to receive Rs 230 crores, Rs 240 crores and Rs 250 crores than the proceeds necessary amouting to Rs 215 crores, Rs 230 crores and Rs 245 crores, they will provide increasing employment to 20 lakh, 25 lakh and 30 lakh workers respectively. But when the proceeds necessary and proceeds expected equal Rs 260 crores the level of employment rises to 35 lakhs. This is the point of effective demand. If we

assume the level of full employment to be 40 lakh workers in the economy, it will necessitate the drawing up of a new aggegate demand price schedule as shown in Table III last column. As a result, the new point of effective demand is 40 lakh workers because both the aggregate demand price and the aggregate supply price equal Rs 275 crores. Beyond this point there is no change in the level of employment which is steady at 40 lakh workers.

Table III: Schedule of Aggregate Demand and Aggregate Supply Prices

Level of Employment(N) (in lakhs)	*Aggregate supply Price(Z) (Rs crores)*	*Aggregate Demand Price(D) (Rs crores)*	
		Old	New
20	215	230	235
25	230	240	245
30	245	250	255
35	260	260	265
40	275	270	275
40	290	280	285
40	305	290	295

Figure 1 illustrates the determination of effective demand where *AD* is the aggregate demand function and *AS* the aggregate supply function. The horizontal axis measures the level of employment in the economy and the vertical axis, the proceeds expected(revenue) and the proceeds necessary (costs). The two curves *AD* and *AS* intersect each other at point *E*. This is effective demand where *ON* workers are employed. At this point, the entrepreneurs' expectations of

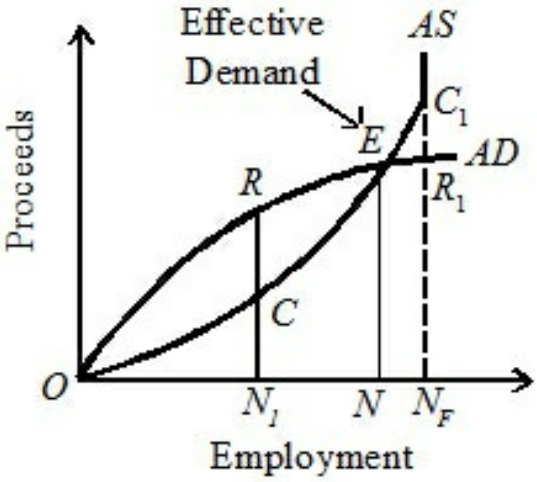


Fig. 1.

profits are maximised. At any point other than this, the entrepreneurs will either incur losses or earn subnormal profits. At ON_1 level of employment, the proceeds expected(revenue) are more than the proceeds necessary (costs), *i.e.*, $RN_1 > CN_1$. This indicates that it is profitable for the entrepreneurs to provide increasing employment to workers till *ON* level is reached where the proceeds expected and necessary equal at point *E*. It would not be, however, profitable for the entrepreneurs to increase employment beyond this to N_F level because the proceeds necessary(costs) exceed the proceeds expected(revenue), *i.e.*, $C_1N_F > R_1N_F$ and they incur losses. Thus E, the point of effective demand, determines the actual level of employment in the economy which is of underemployment equilibrium.

Of the two determinants of effective demand, Keynes regards the aggregate supply function to be given because it depends on the technical conditions of production, the availability of raw materials, machines etc. which do not change in the short run. It is, therefore, the aggregate demand function which plays a vital role in determining the level of employment in the economy. According to Keynes, the aggregate demand function depends on the consumption function and investment function. The cause of unemployment may be a fall in either consumption expenditure or investment expenditure, or both. The level of employment can be raised by increasing either consumption expenditure or investment expenditure, or both. Thus, it is the aggegate demand function which is the "effective" element in the principle of effective demand. Prof. Dillard regards this as the core of the principle of effective demand.

It follows that to raise the economy to the level of full employment requires the raising of the point of effective demand by increasing the aggregate demand. This is illustrated in Figure 2, where *E* is the point of effective demand which determines *ON* level of employment. If ON_F is the level of full employment for the economy, it requires the raising of the point of effective demand. This is possible by raising the aggregate demand curve to AD_1 (last column of Table III) where it intersects aggregate supply curve *AS* at E_1. This is the new point of effective demand which provides an *optimum* level of employment ON_F to the

economy. If the aggregate demand function is raised beyond this point the economy will experience inflation because all the existing resources are fully employed and their supply cannot be increased during the short run, as is apparent from the vertical portion of the *AS* curve in Figure 2.

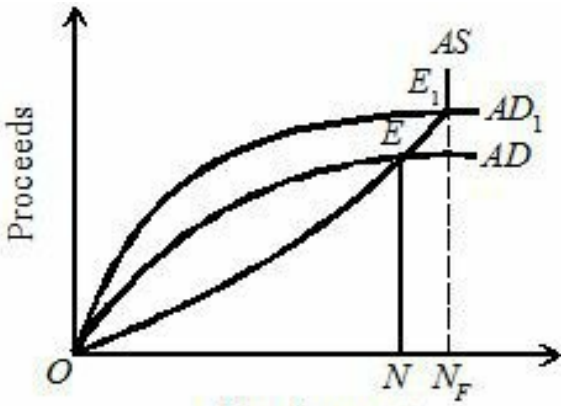


Fig. 2.

IMPORTANCE OF EFFECTIVE DEMAND

The principle of effective demand is the most important contribution of Keynes. It is the soul of the Keynesian theory of employment. Dr Klein attributes the Keynesian revolution solely to the development of a theory of effective demand.

1. Determinant of Employment. Effective demand determines the level of employment in the economy. When effective demand increases, employment also increases, and a decline in effective demand decreases the level of employment. Thus unemployment is caused by a deficiency of effective demand. Effective demand represents the total expenditure on the total output produced at an equilibrium level of employment. It indicates the value of total output which equals national income. National income equals national expenditure. National expenditure consists of expenditure on consumption goods and investment goods. Thus the main determinants of effective demand and the level of employment are consumption and investment. In brief, Effective Demand =Value of National Output=Volume of Employment=National Income=National Expenditure=Expenditure on consumption goods+Expenditure on investment goods.

In the Keynesian analysis of effective demand, consumption and investment expenditures relate to the private sector because Keynes considers government expenditure as autonomous. But the post-Keynesian economists include government expenditure as a component of effective demand. Thus, Effective Demand (*D*) = Private consumption expenditure (*C*) + Private investment (*I*)+Government expenditure (*G*) on both.

We may conclude that the importance of the principle of effective demand lies in pointing out the cause and remedy of unemployment. Unemployment is caused by a deficiency of effective demand and it can be removed by an increase in consumption expenditure or/and investment expenditure and in case the private expenditures are insufficient and ineffective in bringing about the required level of employment, the same can be achieved by government expenditure. Thus the principle of effective demand is the basis of the theory of employment.

2. Repudiation of Say's Law and Full Employment Thesis. The principle of effective demand repudiates Say's law of markets that supply creates its own demand and that full employment equilibrium is a normal situation in the economy. This principle points out that underemployment equilibrium is a normal situation and full employment equilibrium is accidental. In a capitalist economy, supply fails to create its own demand because the whole of the earned income is not spent on the consumption of goods and services. Moreover, the decisions to save and invest are made by different people. As a result, the existence of full employment is not a possibility and the point of effective demand at any time represents underemployment equilibrium.

The *Pigovian view* that full employment can be achieved by a reduction in money wage-cut is also repudiated by this principle. A money wage-cut will bring about a reduction in expenditure on goods and services leading to a fall in effective demand and hence in the level of employment. Thus the importance of this principle lies in repudiating Say's Law and the classical thesis of full employment equilibrium.

3. Role of Investment. The principle of effective demand highlights the significant role of investment in determining the level of employment in the economy. The two determinants of effective demand are consumption and investment expenditures. When income increases consumption expenditure also increases but by less than the increase in income. Thus there arises a gap between income and consumption which leads to decline in the volume of employment. This gap can be bridged by an increase in either consumption expenditure or investment expenditure in order to achieve full employment level of effective demand in the

economy. Since the propensity to consume is stable during the short run, it is not possible to raise the consumption expenditure. Therefore, the level of effective demand and hence of employment can be raised by an increase in investment. In this lies the importance of investment.

4. The Paradox of Poverty in the midst of Potential Plenty. The importance of effective demand lies in explaining the paradox of poverty in the midst of potential plenty in modern capitalism. Effective demand is mainly determined by the aggregate demand function which is composed of consumption expenditure and investment expenditure. A fundamental principle is that when income increases consumption also increases but less than proportionately (*i.e.*, the marginal propensity to consume is less than one). This creates a gap between income and consumption which must be filled up by the required investment expenditure. If the appropriate investment is not forthcoming to fill this gap, it leads to a deficiency of effective demand resulting in unemployment.

It follows that in a poor community, the gap between income and consumption is small because the marginal propensity to consume is high. It will, therefore, have little difficulty in employing all its resources by filling the gap through small investment expenditure.

On the contrary, in a wealthy community the gap between income and consumption is very large because the marginal propensity to consume is low. It will, therefore, require large investment expenditure to fill the gap between income and consumption in order to maintain a high level of income and employment. But in a rich community investment demand is not adequate to fill this gap and there emerges a deficiency of aggregate demand resulting in widespread unemployment. When the aggregate demand falls, the potential wealthy community will be forced to reduce its actual output until it becomes so poor that the excess of output over consumption will be reduced to the actual amount of investment. Further, in such a community there is an accumulated stock of capital assets which weakens the inducement to invest because every new investment competes with an already existing large supply of old capital assets. This inadequacy of investment demand reacts in a cumulative manner on the demand for consumption and will, in turn, lead to a further fall in

employment, output and income. Thus as Keynes said, "The richer the community, the more obvious and outrageous the defects of the economic system that lead to unemployment on a mass scale in the midst of potential plenty because of the deficiency of effective demand."

SELECTED READINGS

W.C., Peterson, *Income, Employment and Growth*, Ch.5

N.F. Keiser, *Macroeconomics*, 1971, Ch.19.

Edward Shapiro, *Macroeconomic Analysis*, 4/e, 1978, Ch.13

EXERCISES

1. What do you mean by Effective Demand ? How is it determined? Discuss its importance.
2. "The logical starting point of Keynes's theory of employment is the principle of effective demand." Discuss.
3. "The intersection of the aggregate supply and demand function determines the equilibrium level of output and employment." Discuss.

CHAPTER 9

THE CONSUMPTION FUNCTION

INTRODUCTION

One of the important tools of the Keynesian economics is the consumption function. This chapter deals with the consumption function, its technical attributes, its importance and its subjective and objective determinants along with Keynes's Psychological Law of Consumption.

MEANING OF CONSUMPTION FUNCTION

The consumption function or propensity to consume refers to income-consumption relationship. It is a "functional relationship between two aggregates, *i.e.*, total consumption and gross national income."[1] Symbolically, the relationship is represented as $C = f(Y)$, where C is consumption, Y is income, and f is the functional relationship. Thus the consumption function indicates a functional relationship between C and Y, where C is the dependent by Y is the independent variable, *i.e.*, C is determined by Y. This relationship is based on the *ceteris paribus* (other things being equal) assumption, as such only income-consumption relationship is considered and all possible influences on consumption are held constant.

1. K.K. Kurihara, *Monetary Theory and Public Policy*, p.152.

In fact, the propensity to consume or consumption function is a schedule of the various amounts of consumption expenditure corresponding to

different levels of income. A hypothetical consumption schedule is given in Table I.

TABLE I : CONSUMPTION SCHEDULE

Income *(Rs Crores)*	Consumption
(*Y*)	*C* = *f* (*Y*)
0	20
60	70
120	120
180	170
240	220
300	270
360	320

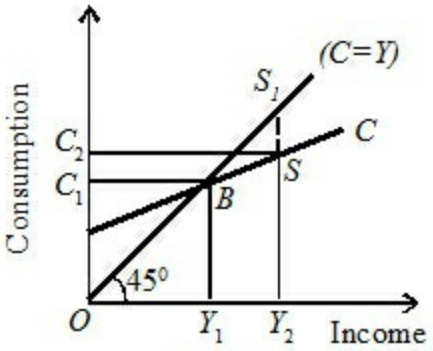


Fig. 1.

Table I shows that consumption is an increasing function of income because consumption expenditure increases with increase in income. Here it is shown that when income is zero during the depression, people spend out of their past savings on consumption because they must eat in order to live. When income is generated in the economy to the extent of Rs 60 crores, it is not sufficient to meet the consumption expenditure of the community so that the consumption expenditure of Rs 70 crores is still above the income amounting to Rs 60 crores (Rs 10 crores are dis-saved). When both consumption expenditure and income equal Rs 120 crores, it is the basic consumption level. After this, income is shown to increase by 60 crores and consumption by 50 crores. This implies a stable consumption function during the short-run as assumed by Keynes. Figure 1 illustrates the consumption function diagrammatically. In the diagram, income is measured horizontally and consumption is measured vertically. 45° is the unity-line where at all levels income and consumption are equal. The C curve is a linear consumption function based on the assumption that consumption

changes by the same amount (Rs 50 crores). Its upward slope to the right indicates that consumption is an increasing function of income. *B* is the break-even point where *C*=*Y* or $OY_1 = OC_1$. When income rises to OY_1

consumption also increases to OC_2, but the increase in consumption is less than the increase in income, $C_1C_2 < Y_1Y_2$. The portion of income not consumed is saved as shown by the vertical distance between 45° line and *C* curve, *i.e.*, SS_1. "Thus the consumption function measures not only the amount spent on consumption but also the amount saved. This is because the propensity to save is merely the propensity not to consume. The 45° line may therefore be regarded as a zero-saving line, and the shape and position of the *C* curve indicate the division of income between consumption and saving."

PROPERTIES OR TECHNICAL ATTRIBUTES OF THE CONSUMPTION FUNCTION

The consumption function has two technical attributes or properties: (i) the average propensity to consume, and (ii) the marginal propensity to consume.

(1) The Average propensity to Consume. "The average propensity to consume may be defined as the ratio of consumption expenditure to any particular level of income."[2] It is found by dividing consumption expenditure by income, or $APC = C/Y$. It is expressed as the percentage or proportion of income consumed. The *APC* at various income levels is shown in column 3 of Table II. The *APC* declines as income increases because the proportion of income spent on consumption decreases. But reverse is the case with *APS* (average propensity to save) which increases with increase in income (see column 4). Thus the *APC* also tells us about the the average propensity to save, $APS=1—APC$.

Diagrammatically, the average propensity to consume is any one point on the *C* curve. In Figure 2 Panel (A), point *R* measures the *APC* of the *C* curve which is OC_1/OY_1. The flattening of the *C* curve to the right shows declining *APC*.

(2) The Marginal Propensity to Consume. "The marginal propensity to consume may be defined as the ratio of the change in consumption to the change in income or as the rate of change in the average propensity to

consume as income changes."[3] It can be found by dividing change in consumption by a change in income, or $MPC = \Delta C/\Delta Y$. The *MPC* is constant at all levels of income as shown in column 5 of Table II. It is 0.83 or 83 per cent because the ratio of change in cosumption to change in income is $\Delta C/\Delta Y = 50/60$. The marginal propensity to save can be derived from the *MPC* by the formula 1–*MPC*. It is 0.17 in our example (see column 6).

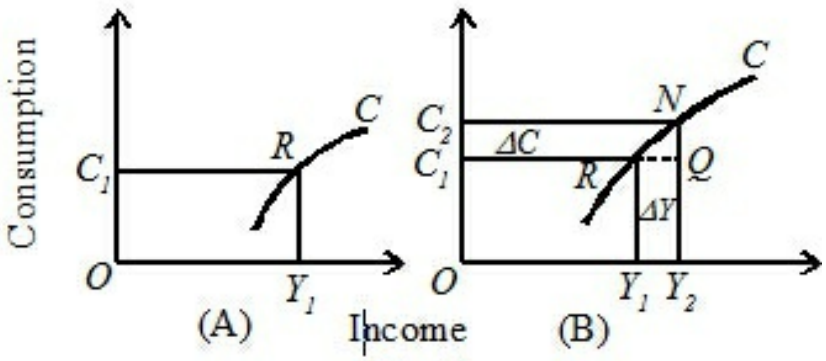


Fig. 2.

Diagrammatically, the marginal propensity to consume is measured by the gradient or slope of the *C* curve. This is shown in Panel (B) by *NQ*/*RQ* where *NQ* is change in consumption (ΔC) and *RQ* is change in income (ΔY) or C_1C_2 / Y_1Y_2.

Significance of MPC

The *MPC* is the rate of change in the *APC*. When income increases, the *MPC* falls but more than the *APC*. Contrariwise, when income falls, the *MPC* rises and the *APC* also rises but at a slower rate than the former. Such changes are only possible during cyclical fluctuations whereas in the short-run there is change in the *MPC* and *MPC<APC*.

TABLE II

(Rs. Crores)

(1) Income Y	(2) Consump-tion(C)	(3) APC=C/Y	(4) APS=S/Y (1-APC)	(5) MPC= ΔC/ΔY	(6) MPS=ΔS/ΔY (1-MPC)
120	120	$\frac{120}{120}$ = 1 or 100%	0	—	—
180	170	$\frac{170}{180}$ = 0.92 or 0.92%	0.08	$\frac{50}{60}$ = 0.83	0.17
240	220	$\frac{220}{240}$ = 0.91 or 91%	0.09	$\frac{50}{60}$ = 0.83	0.17
300	270	$\frac{270}{300}$ = 0.90 or 90%	0.10	$\frac{50}{60}$ = 0.82	0.17
360	320	$\frac{320}{360}$ = 0.88 or 88%	0.12	$\frac{50}{60}$ = 0.83 or 83%	0.17

Keynes is concerned primarily with the *MPC*, for his analysis pertains to the shor-run while the *APC* is useful in the long-run analysis. The post-Keynesian economists have come to the conclusion that over the long-run *APC* and *MPC* are equal and approximate 0.9. In the Keynesian analysis the *MPC* is given more prominence. Its value is assumed to be positive and less than unity which means that when income increases the whole of it is not spent on consumption. On the contrary, when income falls, consumption expenditure does not decline in the same proportion and never becomes zero. The Keynesian hypothesis that the marginal propensity to consume is positive but less than unity($O < \Delta C/\Delta Y < 1$) is of great analytical and practical significance. Besides telling us that

consumption is an increasing function of income and it increases by less than the increment of income, this hypothesis helps in explaining "(*a*) the theoretical possibility of general over production or 'underemployment equilibrium,' and also (*b*) the relative stability of a highly developed industrial economy. For it is implied that the gap between income and consumption at all high levels of income is too wide to be easily filled by investment with the possible consequence that the economy may fluctuate around an underemployment equilibrium."[4] Thus the economic significance of the *MPC* lies in filling the gap between income and consumption through planned investment to maintain the desired level of income. Further, its importance lies in the multiplier theory. Th higher the *MPC*, the higher the multiplier and vice versa. The *MPC* is low in the case of the rich people and high in the case of the poor. This accounts for high *MPC* in underdeveloped countries and low in advanced countries.

---

2. K.K. Kurihara, *Introduction to Keynesian Dynamics*, p. 32.

3. K.K. Kurihara, *Monetary Theory and Public Policy*, p.156.

4. K.K. Kurihara, *Introduction to Keynesian Dynamics*, p.36.

KEYNES'S PSYCHOLOGICAL LAW OF CONSUMPTION

Keynes propounded the fundamental psychological law of consumption which forms the basis of the consumption function. He wrote, "The fundamental psychological law upon which we are entitled to depend with great confidence both *a prior* from our knowledge of human nature and from the detailed facts of experience, is that men are disposed as a rule and on the average to increase their consumption as their income increases but not by as much as the increase in their income." The law implies that there is a tendency on the part of the people to spend on consumption less than the full increment of income.

Propositions of the Law

This law has three related propositions:

(1) When income increases, consumption expenditure also increases but

by a smaller amount. The reason is that as income increases, our wants are satisfied side by side, so that the need to spend more on consumer goods diminishes. It does not mean that the consumption expenditure falls with the increase in income. In fact, the consumption expenditure increases with increase in income but less than proportionately.

(2) The increased income will be divided in some proportion between consumption expenditure and saving. This follows from the above proposition because when the whole of increased income is not spent on consumption, the remaining is saved. In this way, consumption and saving move together.

(3) Increase in income always leads to an increase in both consumption and saving. This means that increased income is unlikely to lead either to fall in consumption or saving than before. This is based on the above propositions because as income increases consumption also increases but by a smaller amount than before which leads to an increase in saving. Thus with increased income both consumption and saving increase.

The three propositions of the law can be explained with the help of the following Table III.

TABLE III

(Rs Crores)

Income (Y)	*Consumption (C)*	*Savings (S=Y—C)*
0	20	—20
60	70	—10
120	120	0
180	170	10
240	220	20
300	270	30
360	320	40

Proposition (1) : Income increases by Rs 60 crores and the increase in consumption is by Rs 50 crores. The consumption expenditure is, however, increasing with increase in income, *i.e.*, Rs 170, 220, 270 and

320 crores against Rs 180, 240, 300 and 360 crores respectively.

Proposition (2) : The increased income of Rs 60 crores in each case is divided in some proportion between consumption and saving(*i.e.*, Rs 50 crores and Rs 10 crores).

Proposition (3) : As income increases from Rs 120 to 180, 240, 300 and 360 crores, consumption also increases from Rs 120 to 170, 220, 270, 320 crores, along with increase in saving from Rs 0 to 10, 20, 30 and 40 crores respectively. With increase in income neither consumption nor saving have fallen.

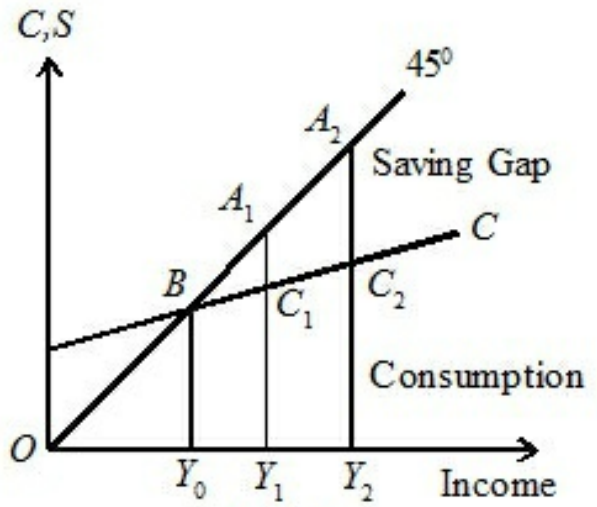


Fig. 3.

Diagrammatically, the three propositions are explained in Figure 3. Here, income is measured horizontally and consumption and saving are measured on the vertical axis. *C* is the consumption function curve and 45° line represents income. *Proposition* (1): When income increases from OY_0 to OY_1 consumption also increases from BY_0 to C_1Y_1 but the increase in consumption is less than the increase in income, *i.e.*, $C_1Y_1 < A_1Y_1$ ($=OY_1$) by A_1C_1. *Proposition* (2): When income increases to OY_1 and OY_2, it is divided in some proportion between consumption C_1Y_1 and C_2Y_2 and saving A_1C_1 and A_2C_2 respectively. *Proposition* (3): Increases in income to OY_1 and OY_2 lead to increased consumption $C_2Y_2 > C_1Y_1$ and increased saving $A_2C_2 > A_1C_1$ than before. It is clear from the widening area below the *C* curve and saving gap between 45° line and *C* curve.

Its Assumptions

Keynes's Law is based on the following assumptions:

1. It assumes a Constant Psychological and Institutional Complex. This law is based on the assumption that the psychological and institutional complexes influencing consumption expenditure remain constant. Such complexes are income distribution, tastes, habits, social customs, price movements, population growth, etc. In the short run, they do not change and consumption depends on income alone. The constancy of these complexes is the fundamental cause of the stable consumption function.

2. It assumes the Existence of Normal Conditions. The law holds good under normal conditions. If, however, the economy is faced with abnormal and extraordinary circumstances like war, revolution or hyperinflation, the law will not operate. People may spend the whole of increased income on consumption.

3. It assumes the Existence of a Laissez-faire Capitalist Economy. The law operates in a rich capitalist economy where there is no government intervention. People should be free to spend increased income. In the case of regulation of private enterprise and consumption expenditures by the state, the law breaks down. Thus the law is inoperative in socialist or state controlled and regulated economies.

Professor Kurihara opines that "Keynes's law based on these assumptions may be regarded as a rough approximation to the actual macro-behaviour of free consumers in the normal short period."[6]

IMPLICATIONS OF KEYNES'S LAW (OR IMPORTANCE OF THE CONSUMPTION FUNCTION)

Keynes's psychological law has important implications which in fact point towards the importance of the consumption function[7] because the latter is based on the former. The following are its implications:

1. Invalidates Say's Law. Say's Law states that supply creates its own demand. Therefore, there cannot be general overproduction or general unemployment. Keynes's psychological law invalidates Say's Law because as income increases, consumption also increases but by a smaller amount.

In other words, all that is produced (income) is not taken off the market (spent), as income increases. Thus supply fails to create its own demand. Rather it exceeds demand and leads to general overproduction and glut of commodities in the market. As a result, producers stop production and there is mass unemployment.

2. Need for State Intervention. As a corollary to the above, the psychological law highlights the need for state intervention. Say's Law is based on the existence of *laissez-faire* policy and its refutation implies that the economic system is not self-adjusting. So when consumption does not increase by the full increment of income and consequently there is general overproduction and mass unemployment, the necessity of state intervention arises in the economy to avert general overproduction and unemployment through public policy.

3. Crucial Importance of Investment. Keynes's psychological law stresses the vital point that people fail to spend on consumption the full increment of income. This tendency creates a gap between income and consumption which can only be filled by either increaed investment or consumption. If either of them fail to rise, output and employment will inevitably fall. Since the consumption function is stable in the short-run, the gap between income and consumption can only be filled by an increase in investment. Thus the psychological law emphasises the crucial role of investment in Keynes's theory. It is the inadequacy of investment which results in unemployment and logically, the remedy to overcome unemployment is increase in investment.

4. Existence of Underemployment Equilibrium. Keynes's notion of underemployment equilibrium is also based on the psychological law of consumption. The point of effective demand which determines the equilibrium level of employment is not of full employment but of underemployment because consumers do not spend the full increment of their income on consumption and there remains a deficiency in aggregate demand. Full employment equilibrium level can, however, be reached if the state increases investment to match the gap between income and consumption.

5. Declining Tendency of the Marginal Efficiency of Capital. The

psychological law also points towards the tendency of declining marginal efficiency of capital in a *laissez-faire* economy. When income increases and consumption does not increase to the same extent, there is a fall in demand for consumer goods. This results in glut of commondities in the market. The producers will reduce production which will, in turn, bring a decline in the demand for capital goods and hence in the expected rate of profit and business expectations. It implies a decline in the marginal efficiency of capital. It is not possible to arrest this process of declining tendency of marginal efficiency of capital unless th propensity to consume rises. But such a possibility can exist only in the long run when the psychological law of consumption does not hold good.

---

6. K.K. Kurihara, *Monetary Theory and Public Policy*, p.160.

7. While dealing with the importance of the Consumption Function, the students are advised to delete word 'psychological law' and instead write the word 'Consumption Function'.

6. Danger of Permanent Over-saving or Under-investment Gap. Keynes's psychological law points out that there is always a danger of an over-saving or under-investment gap appearing in the capitalist economy because as people become rich the gap between income and consumption widens. This long-run tendency of increase in saving and fall in investment is characterised as secular stagnation. When people are rich, their propensity to consume is low and they save more. This implies low demand which leads to decline in investment. Thus the tendency is for secular stagnation in the economy.

7. Unique Nature of Income Propagation. The fact that the entire increased income is not spent on consumption explains the multiplier theory. The multiplier theory or the process of income propagation tells that when an initial injection of investment is made in the economy, it leads to smaller successive increments of income. This is due to the fact that people do not spend their full increment of income on consumption. In fact, the value of multiplier is derived from the marginal propensity to consume, *i.e.*, Multiplier = 1—1/*MPC*. The higher the *MPC*, the higher the value of the multiplier, and vice versa.

8. Explanation of the Turning Points of the Business Cycles. This law

explains the turning points of a business cycle. Before the economy reaches the full employment level, the downturn starts because people fail to spend the full increment of their income on consumption. This leads to fall in demand, overproduction, unemployment and decline in the marginal efficiency of capital. Panel (A) of Figure 4 shows this downturn movement. When income increases above the breakeven point by Y_1Y_2, consumption expenditure increases by a smaller amount C_1C_2, $C_1C_2 < Y_1Y_2$. Before the economy reaches the full-employment income level Y_F , the downturn will start because the gap between 45° line and *C* curve continues to widen.

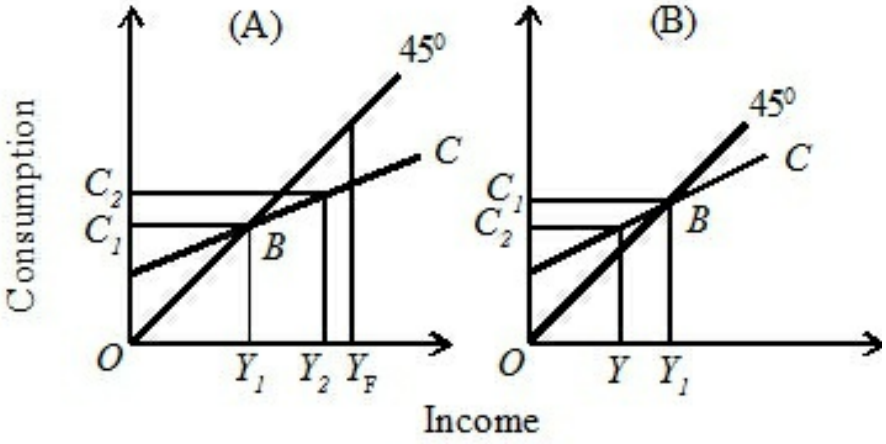


Fig. 4.

Conversely, the upturn in the economy starts before it reaches the stage of complete depression because when income falls, consumption also falls but by less than the fall in income. People continue to buy consumer goods even when their income falls. So when the excess stock of commodities is exhausted in the community during a depression, the existence of consumer expenditure on goods leads to revival. This is best explained with the help of Panel (B) where below the breakeven point *B*, the *C* curve is above the 45° income line. The fact that the consumption function curve *C* is above the income line shows that revival will start before income falls to zero. This is because the fall in consumption C_1C_2 is less than the fall in income Y_1Y .

DETERMINANTS OF THE CONSUMPTION FUNCTION

Keynes mentions two principal factors which influence the consumption

function and determine its slope and position. They are (i) the subjective factors, and (ii) the objective factors.

The subjective factors are endogenous or internal to the economic system. They include psychological characteristics of human nature, social practices and institutions and social arrangements. They "are unlikely to undergo a material change over a short period of time except in abnormal or revolutionary circumstances." They, therefore, determine the slope and position of the *C* curve which is fairly stable in the short-run.

The objective factors are exogenous or external to the economic system. They may, therefore, undergo rapid changes and may cause marked shifts in the consumption function (*i.e.*, the C curve).

The subjective and objective factors are discussed below.

Subjective Factors

Keynes's subjective factors basically underlie and determine the form (*i.e.*, slope and position) of the consumption function. As already noted above, the subjective factors are the psychological characteristics of human nature, social practices and institutions, especially the behaviour patterns of business concerns with respect to wage and dividend payments and retained earnings, and social arrangements affecting the distribution of income. There are two motives of subjective factors: individual and business.

1. Individual Motives. First, there are eight motives "which lead individuals to refrain from spending out of their incomes." They are :

(i) the desire to build reserves for unforeseen contingencies;

(ii) the desire to provide for anticipated future needs, *i.e.*, old age, sickness, etc.;

(iii) the desire to enjoy and enlarged future income by way of interest and appreciation;

(iv) the desire to enjoy a gradually increasing expenditure in order to improve the standard of living;

(v) the desire to enjoy a sense of independence and power to do things;

(vi) the desire to secure a "*masse de manoeuver*" to carry out speculative or business projects;

(vii) the desire to bequeath a fortune;

(viii) the desire to satisfy a pure miserly instinct.

2. Business Motives. The subjective factors are also influenced by the behaviour of business corporations and governments. Keynes lists four motives for accumulation on their part:

(i) *enterprise*, the desire to do big things and to expand;

(ii) *liquidity*, the desire to meet emergencies and difficulties successfully;

(iii) *income raise*, the desire to secure large income and to show successful management;

(iv) *financial prudence*, the desire to provide adequate financial resources against depreciation and obsolescence, and to discharge debt.

These factors remain constant during the short-run and keep the consumption fuction stable.

Objective Factors

The following objective factors are given by Keynes.

1. Changes in the Wage Level. If the wage rate rises, the consumption function shifts upward. The workers having a high propensity to consume spend more out of their increased income and this tends to shift the *C* curve upward. If, however, the rise in the wage rate is accompanied by a more than proportionate rise in the price level, the real wage rate will fall and it will tend to shift the *C* curve downward. A cut in the wage rate will

also reduce the consumption function of the community due to a fall in income, employmnet and output. This will shift the curve downward.

2. Windfall Gains or Losses. Unexpected changes in the stock market leading to gains or losses tend to shift the consumption function upward or downward. For instance, the phenomenal windfall gains due to the stock-market boom in the American economy after 1925 led to a rise in the consumption spending of the stock-holders by roughly in proportion to the increased income and as a result the consumption function shifted upward. Similarly, unexpected losses in the stock market lead to the downward shifting of the *C* curve.

3. Changes in the Fiscal Policy. Changes in fiscal policy in the form of taxation and public expenditure affect the consumption function. Heavy commodity taxation adversely affect the consumption function by reducing the disposable income of the people. This is what actually happened during the Second World War when the consumption function shifted downward due to heavy indirect taxation, rationing and price controls. On the other hand, the policy of progressive taxation along with that of public expenditure on welfare programmes tends to shift the consumption function upward by altering the distribution of income.

4. Changes in Expectations. Changes in future expectations also affect the propensity to consume. If a war is expected in the near future, people start hoarding durable and semi-durable commodities in anticipation of future scarcity and rising prices. As a result, people buy much in excess of their current needs and the consumption function shifts upward. On the contrary, if it is expected that prices are likely to fall in the future, people would buy only those things which are very essential. It will lead to a fall in consumption demand and to a downward shift of the consumption function.

5. Changes in the Rate of Interest. Substantial changes in the market rate of interest may influence the consumption function indirectly. There are several ways in which the rate of interest may affect the consumption function. A rise in rate of interest will lead to a fall in the price of bonds, thereby tending to discourage the propensity to consume of the bond-holders.

It may also have the effect of substituting one type of assets for another. People may be encouraged to save rather than invest in bonds. In case they are buying durable consumer goods like refrigerators, scooters, etc. on hire-purchase system they will tend to postpone their purchases when the rate of interest rises. They will have to pay more in instalments and thus their consumption function will shift downward. Keynes wrote, 'Over a long period, substantial changes in the rate of interest probably tend to modify social habits considerably."

Besides, these five factors, Keynes also listed changes in accounting practice with respect to depreciation. This factor has been rejected by Hansen who opines that "it is not a factor which can be thought to change violently in the short-run and it was a mistake for Keynes to include it here." However, we add some of the other objective factors listed by Keynes's followers.

6. Financial Policies of Corporations—Financial policies of corporations with regard to income retention, dividend payments and reinvestments tend to affect the consumption function in several ways. If corporations keep more money in the form of reserves, dividend payments to shareholders will be less, this will have the effect of reducing the income of the shareholders and the consumption function will shift downward.

7. Holding of Liquid Assets. The amount of liquid assets in the form of cash balances, savings and government bonds in the hands of consumers also influence the consumption function. If people hold larger liquid assets they will have a tendency to spend more out of their current income and the propensity to consume will move upward, and vice versa. Pigou was of the view that with a cut in money wage, prices fall and the real value of such assets increases. This tends to shift the consumption function upward. This is called the "Pigou Effect."

8. The Distribution of Income. The distribution of income in the community also determines the shape of the consumption function. If there are large disparities in income distribution between the rich and the poor, the consumption function is low because the rich have a low propensity to consume and the poor with a very low income are unable to

spend more on consumption. If through progressive taxation and other fiscal measures, the inequalities of income and wealth are reduced, the consumption function will shift upward because with the increase in the income of the poor their consumption expenditure will increase more than the reduction in the expenditure of the rich. "Moreover, if the distribution of income is significantly altered for political or humanitarian reasons, consumer habits themselves may undergo such changes as to cause the position or shape of the entire consumption function to vary perceptibly."

9. Attitude toward Saving. The consumption function is also influenced by people's attitude toward saving. If they value future consumption more than present consumption, they will tend to save more and the consumption function will shift downward. This tendency may be reinforced by the state through compulsory life insurance, provident fund and other social insurance schemes to keep the consumption function low. In a high-saving economy, the consumption function is low.

10. Duesenberry Hypothesis. James Duesenberry has propounded a relative income hypothesis affecting the consumption function. The first part of this hypothesis relates to the 'demonstration effect.' There is a tendency in human beings not only to keep up with the Joneses but also to surpass the Joneses, that is, the tendency is to strive constantly toward a higher consumption level and to emulate the consumption patterns of one's rich neighbours and even to surpass them. Thus consumption preferences are interdependent. The second part is the 'past peak of income' hypothesis which explains the short-run fluctuations in consumption. Once the community reaches a particular income level and standard of living, it is reluctant to come down to a lower level of consumption during a recession. Consumption is sustained by the reduction in current saving and vice versa. So there is no shift in the consumption function during the short-run. There is simply an upward-downward movement on the same consumption function when income rises or falls during the short-run.

We may conclude with Professor Hansen "that except for quite abnormal or revolutionary changes in certain *objective* factors...shifts in the 'propensity to consume out of a given income' are not likely to be of more

than secondary importance."

MEASURES TO RAISE THE PROPENSITY TO CONSUME

The propensity to consume remains stable during the short-run due to the existence of certain psychological and institutional factors in the society. But "employment can only increase *pari passu* with an increase in investment; unless, indeed, there is a change in the propensity to consume," as pointed out by Keynes. Therefore, it is significant to study the measures which tend to raise the propensity to consume.

1. Income Redistribution. Redistribution of income in favour of the poor tends to raise the propensity to consume because the marginal propensity to consume of the low income groups is high in comparison to the rich. Therefore, the propensity to consume can be raised by transferring income and wealth from the rich to the poor. This can be done by the state through its taxation and public spending policies. By imposing progressive taxes on incomes, expenditures, estates, capital gains, etc., the state is able to mobilise larger revenues for providing more facilities to the poor. But care should be taken that such taxation should not adversely affect investment. Secondly, the state can increase the income of the poor through a judicious public expenditure programme. By starting public works, it is in a position to increase income by providing larger employment opportunities to the unemployed. The provision for free education., free mid-day meals, free health services, low-rent housing, etc. indirectly helps in increasing the income of the workers and tends to raise their consumption expenditure. Such social expenditures by the state also increase the efficiency of the workers which, in turn, leads to a rise in their wages.

2. Increased Wages. If wages are raised, they will have a direct effect in shifting the consumption fuction upward. But a policy of high wages adversely affects the level of employment in the economy for it is not possible to raise the marginal revenue productivity of labour in the short-run. If wages are raised in such a situation, costs will rise in the absence of increase in the marginal revenue productivity of labour and the economy is likely to experience unemployment. Therefore, the long-run wage

policy should be such that wages increase *pari passu* with increase in labour productivity. This will tend to raise the level of consumption in the economy.

3. Social Security Measures. Social security measures tend to raise the consumption function in the long-run. Provisions for unemployment relief, medical facilities, old age pension, etc. remove future uncertainties and the tendency to save is reduced on the part of the people. The state should, therefore, provide larger social security measures to raise the propensity to consume of the people. Unemployment relief and old age pensions tend to maintain a high consumption expenditure even during a depression and thus help bring revival in the economy. So social security measures tend to raise the consumption function both in periods of prosperity and depression.

4. Credit Facilities. Cheap and easy credit facilities help in shifting the consumption function upward. When loans are easily and cheaply available to the people, they buy more durable consumer goods like scooters, televisions, refrigerators, etc. This tends to raise the propensity to consume. To purchase these things on instalment basis or on hire-purchase system produces the same effect. Thus credit facilities in various ways help raise the propensity to consume of durable consumer goods.

5. Advertisement. Advertisement is one of the most significant ways to raise the propensity to consume in modern times. Advertisement and propaganda through the various media of radio, television, cinema, newspaper, etc. make the consumers familiar with the uses of products. The consumers are attracted toward them and they tend to buy them. This raises their propensity to consume.

6. Development of the Means of Transport. Well developed means of transport also tend to shift the consumption function upward. The movement of goods from the maufacturing centres to the different parts of the country becomes easy. The size of the market expands. The prices may also fall due to the reduction of transport costs. Things are available to the people in their respective towns. All this has the tendency to raise the consumption function.

7. Urbanisation. As a corollary to the above, urbanisation helps raise the propensity to consume. When urbanisation takes place, people move from the rural to the urban areas. They are enamoured by new articles and inluenced by the demonstration effect. This tends to shift the consumption function upward. Thus the state should follow the policy of deliberate urbanisation for the purpose of raising the consumption function.

SELECTED READINGS

J.M. Keynes, *The General Theory of Employment, Interest and Money*, 1936, Ch.8 and 9.

G. Ackley, *Macroeconomic Theory*, 1961, Ch.X.

F.S. Brooman, *Macro-Economics*, 4/e, 1970, Ch.V.

T.F. Dernberg and D.M. McDougall, *Macroeconomics*, 5/e, 1976, Ch.4.

EXERCISES

1. Explain the concept of propensity to consume and show how it occupies a strategic position in the theory of employment.
2. Explain Keynes's Psychological Law and discuss its implications.
3. What do you mean by consumption function ? Explain the factors influencing consumption function.
4. "Keynes's discovery of the consumption function must be regarded as one of the major breakthroughs of modern economics." Discuss.
5. "Keynes's consumption function is an epoch making tool in economic analysis." Discuss.
6. "Keynes's consumption function is an epoch-making contribution to the tools of economic analysis." Discuss this statement and bring out the importance of consumption function in macro-economic analysis.

CHAPTER 10

THEORY OF THE CONSUMPTION FUNCTION

INTRODUCTION

In the previous chapter, we studied Keynes' consumption function. Economists after Keynes verified his consumption function in a number of empirical studies. Their findings have led to a few new theories or hypotheses of consumption. Before discussing them, we shall first explain Keynes' consumption function theory.

KEYNES' CONSUMPTION FUNCTION : THE ABSOLUTE INCOME HYPOTHESIS

Keynes in his *General Theory* postulated that aggregate consumption is a function of aggregate current disposable income. The relation between consumption and income is based on his Fundamental Psychological Law of Consumption which states that when income increases consumption expenditure also increases but by a smaller amount.

The Keynesian consumption function is written as

$$C = a + cY \qquad a > 0, \qquad 0 < c < 1$$

where a is the intercept, a constant which measures consumption at a zero level of disposal income; c is the marginal propensity to consume (MPC); and Y is the disposal income.

The above relation that consumption is a function of current disposable

income whether linear or non-linear, is called the absolute income hypothesis. This consumption function has the following properties:

1. As income increases, average propensity to consume ($APC = C/Y$) falls.

2. The marginal propensity to consume (MPC) is positive but less than unity ($0 < c < 1$) so that higher income leads to higher consumption.

3. The consumption expenditure increases (or decreases) with increase (or decrease) in income but non-proportionally. This non-proportional consumption function implies that in the short-run average and marginal propensities do not coincide ($APC > MPC$).

4. This consumption function is stable both in the short-run and the long-run.

This consumption function is explained in Fig. 1 where $C = a + cY$ is the consumption function. At point E on the C curve the income level is OY_1. At this point, $APC > MPC$ where $APC = OC_1/OY_1$ and $MPC = \Delta C/\Delta Y = ER/RE_O$. This shows disproportional consumption function. The intercept a shows the level of consumption corresponding to a zero level of income. At income level OY_0 where the curve C intersects the 45° line, point E_0 represents APC $(=OC_0/OY_0)$. Below the income level OY_0, consumption is more than income. In this range, $APC > 1$. Above the income level OY_0, consumption increases less than proportionately with income so that APC declines and it is less than one.

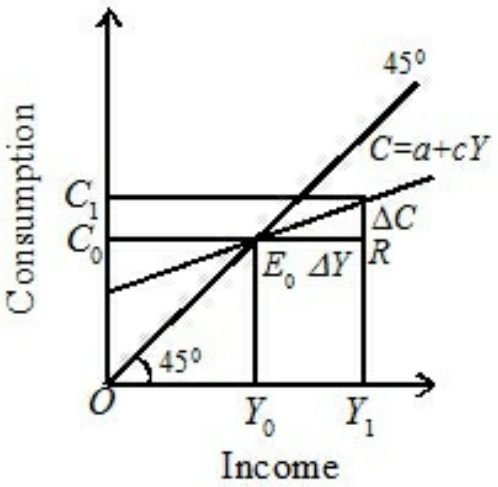


Fig 1

Empirical Studies

Keynes put forth this hypothesis on the basis of "knowledge of human nature" 'and "delailed facts of experience". His followers in a number of

empirical studies based on cross-section budget figures and short-run time series data in the late 1930s and mid-1940s confirmed his hypothesis. They found that families with higher income levels consumed more which confirms that MPC is greater than zero ($c>0$), but by less than the increase in income ($c<1$). They also found that families with higher income levels saved more and so consumed a smaller proportion of income which confirms that APC falls as income rises.

THE CONSUMPTION PUZZLE

Keynes' assertion that the APC falls as income rises led some Keynesians to formulate the secular stagnation thesis around 1940. According to these economists, as incomes grew in the economy, households would save more and consume less. As a result, aggregate demand would fall short of output. If the government spending was not increased at a faster rate than income, the economy would lapse into stagnation. But after World War II, the American economy experienced inflation rather than stagnation even when the government expenditures were reduced below 1941 level in constant dollars. The Keynesian consumption function had been proved wrong. This was due to the conversion of government bonds into liquid assets after the War by the households in order to meet their pent up demand for consumer goods.

In 1946, Kuznets studied the consumption and income data for the United States during the period 1869-1938 and estimated the consumption function for this period as 0.9.[1] Further, he arrived at two conclusions: *one*, over the long-run, on the average, the *APC* did not show any downward trend so that the *MPC* equalled the *APC* as income increased along a long-run trend. This means that the consumption function is a

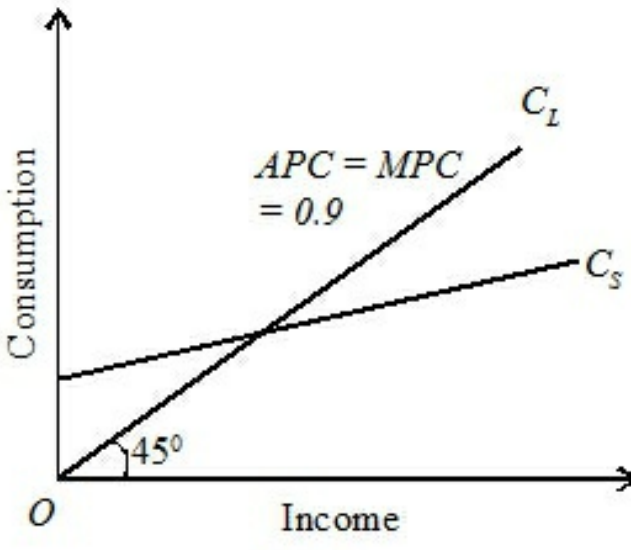


Fig 2

straight line through the origin, as shown by the C_L line in Fig. 2, and *two*, the years in which the *APC* was below the long-run average were boom periods, and the years in which the *APC* was above the long-run average were of slump periods. This implies that in the short-run as income changes over the business cycle, the *MPC* is less than the *APC*, as shown by the *Cs* curve in Fig. 2

These findings were later verified by Goldsmith in 1955 who found the long-run consumption function to be stable at 0.87.[2] Thus these two studies revealed that for the short-run time series, the consumption function is non-proportional because *APC* > *MPC* and for the long-run time series, the consumption function is proportional, *APC* = *MPC*.

The failure of the secular stagnation hypothesis and the findings of Kuznets and Goldsmith were a puzzle to the economists which is known as the *consumption puzzle*. Figure 2 illustrates this puzzle where there are two consumption functions. *Cs* is the Keynesian consumption function which is non-proportionl (*APC* > *MPC*) and based on the short-run time series data. C_L is the long-run proportional consumption function (*APC* = *MPC*) based on long-run time series data. Over the years, economists have been engaged in solving this puzzle by reconciling the two consumption functions.

We study below a few important theories which try to reconcile the two consumption functions.

---

1. Kuznets, *National Product since 1869*, 1946.

2. Goldsmith, *A Study of Saving in the United States*, Vol. I. 1955.

THE DRIFT THEORY OF CONSUMPTION

One of the first attempts to reconcile the short-run and long-run consumption functions was by Arhur Smithies[3] and James Tobin[4]. They tested Keynes' absolute income hypothesis in separate studies and came to the conclusion that the short-run relationship between consumption and income is non-proportional but the time series data show the long-run

relationship to be proportional.

The latter consumption-income behaviour results through an upward shift or "drift" in the short-run non-proportional consumption function due to factors other than income. Smithies and Tobin discuss the following factors:

1. Asset Holdings. Tobin introduced asset holdings in the budget studies of negro and white families to test this hypothesis. He came to the conclusion that the increase in the asset holdings of families tends to increase their propensity to consume thereby leading to an upward shift in their consumption function.

2. New Products. Since the end of the Second World War, a variety of new household consumer goods have come into existence at a rapid rate. The introduction of new products tends to shift the consumption function upward.

3. Urbanisation. Since the post-War period, there has been an increased tendency toward urbanisation. This movement of population from rural to urban areas has tended to shift the consumption function upward because the propensity to consume of the urban wage earners is higher than that of the farm workers.

4. Age Distribution. There has been a continuous increase in the percentage of old people in the total population over the long-run. Though the old people do not earn but they do consume commodities. Consequently, the increase in their numbers has tended to shift the consumption function upward.

5. Decline in Saving Motive. The growth of social security system which makes automatic saving and guarantees income during illness. Unemployment disability and old age has increased the propensity to consume.

6. Consumer Credit. The increasing availability and convenience of short-term consumer credit shifts the consumption function upward. The greater ease of buying consumer goods with credit cards, debit cards, use

of ATMs and cheques, and availability of installment buying causes an upward shift in the consumption function.

7. Expectation of Income Increasing . Average real wages of workers have increased and they expect them to rise in the future. These cause an upward shift in the consumption function. Those who expect higher future earnings tend to reduce their savings or even borrow to increase their present consumption.

The consumption drift theory is explained in Fig. 3 where C_L is the long-run consumption function which shows the proportional relationship between consumption and income as we move along it. C_{S1} and C_{S2} are the short-run consumption functions which cut the long-run consumption function C_L at points A and B. But due to the factors mentioned above, they tend to "drift" upward from point A to point B along the C_L curve. Each point such as A and B on the C_L curve represents an average of all the values of factors included in the corresponding short-run functions, C_{S1} and C_{S2} respectively and the long-run function, C_L, connecting all the average values. But the movement along the dotted portion of the short-run consumption functions, C_{S1} and C_{S2}, would cause consumption not to increase in proportion to the increase in income.

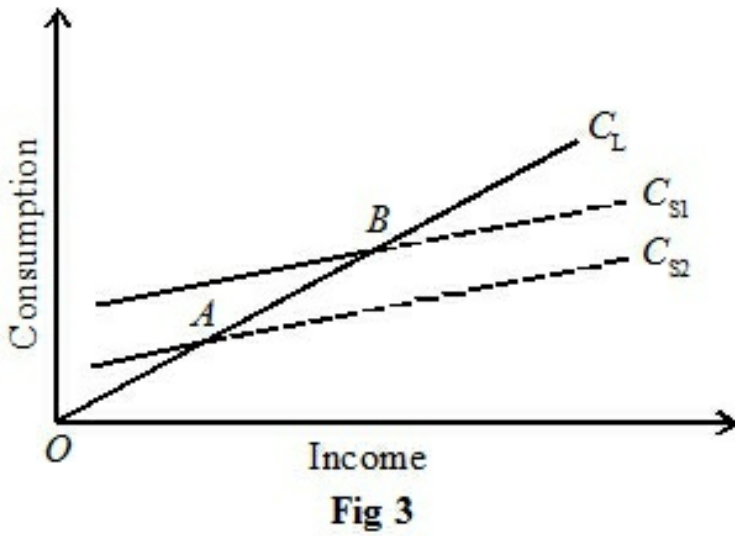


Fig 3

---

3. Arthur Smithies, 'Forecasting Behaviour Demand : I," *Econometrica*, Jan, 1945.

4. James Tobin, "Relative Income, Absolute Income and Saving," in *Money, Trade and Economic Growth : Essays in Honour of John H. Williams*, 1951.

Its Criticisms

The great merit of this theory is that it lays stress on factors other than in income which affect the consumer behaviour. In this sense, it represents a major advance in the theory of the consumption function. However, it has its shortcomings.

1. The theory does not tell the rate of upward drift along the C_L curve. It appears to be a matter of chance.

2. It is just a coincidence if the factors explained above cause the consumption function to increase proportionately with increase in income so that the average of the values in the short-run consumption function equals a fixed proportion of income.

3. According to Duesenberry, all the factors mentioned as causes of the upward shift are not likely to have sufficient force to change the consumption-savings relationship to such an extent as to cause the drift.

4. Duesenberry also points out that many of the factors such as decline in saving motive would lead to a secular fall in the consumption function. Such saving plans as life insurance and pension programs tend to increase savings and decrease the consumption function. Moreover, people want more supplementary savings to meet post-retirement needs which tend to decrease their current consumption*.

---

* As an alternative to the consumption-drift theory, Duesenberry gave his Relative Income or Consumption-Ratchet Hypothesis. Duesenberry was a student of Smithies.

THE RELATIVE INCOME HYPOTHESIS

The relative income hypothesis of James Duesenberry[5] is based on the rejection of the two fundamental assumptions of the consumption theory of Keynes. Duesenberry states that (1) every individual's consumption behaviour is not independent but *interdependent* of the behaviour of every other individual, and (2) that consumption relations are irreversible and not reversible in time.

In formulating his theory of the consumption function, Duesenberry

writes: "A real understanding of the problem of consumer behaviour must begin with a full recongnition of the social character of consumption patterns." By the "social character of consumption paterns" he means the tendency in human beings not only "to keep up with the Joneses" but also to surpass the Joneses. Joneses refers to rich neighbours. In other words, the tendency is to strive constantly toward a higher consumption level and to emulate the consumption patterns of one's rich neighbours and associates. Thus consumers' preferences are interdependent. It is, however, differences in *relative incomes* that determine the consumption expenditures in a community. A rich person will have a lower *APC* because he will need a smaller portion of his income to maintain his consumption pattern. On the other hand, a relatively poor man will have a higher *APC* because he tries to keep up with the consumption standards of his neighbours or associates. This provides the explanation of the constancy of the long-run *APC* because lower and higher *APCs* would balance out in the aggregate. Thus even if the absolute size of income in a country *increases*, the *APC* for the economy as a whole at the higher absolute level of income would be constant. But when income decreases, consumption does not fall in the same proportion because of the *Ratchet Effect*.

The Ratchet Effect

The second part of the Duesenberry theory is the "past peak of income" hypothesis which explains the short-run fluctuations in the consumption function and refutes the Keynesian assumption that consumption relations are reversible. The hypothesis states that during a period of prosperity, consumption will increase and gradually adjust itself to a higher level. Once people reach a particular peak income level and become accustomed to this standard of living, they are not prepared to reduce their consumption pattern during a recession. As income falls, consumption declines but proportionately *less* than the decrease in income because the consumer dissaves to sustain consumption. On the other hand, when income increases during the recovery period, consumption rises gradually with a rapid increase in saving. Economists call this the *Ratchet Effect*.

Duesenberry combines his two related hypothesis in the following form:

$$\frac{C_t}{Y_t} = a - c\frac{Y_t}{Y_o}$$

5. James S. Duesenberry, *Income, Saving and the Theory of Consumer Behaviour*, 1949.

where C and Y are consumption and income respectively, t refers to the current period and the subscript (o) refers to the previous peak, a is a constant relating to the positive autonomous consumption and c is the consumption function. In this equation, the consumption-income ratio in the current period (C_t / Y_t) is regarded as function of Y_t / Y_o, that is, the ratio of current income to the previous peak income. If this ratio is constant, as in periods of steadily rising income, the current consumption income ratio is constant. During recession when current income (Y_t) falls below the previous peak income (Y_O), the current consumption income ratio (C_t / Y_t) will increase.

The relative income hypothesis is explained graphically in Fig. 4 where C_L is the long-run consumption function and C_{S1} and C_{S2} are the short-run consumption functions. Suppose income is at the peak level of OY_1 where E_1Y_1 is consumption . Now income falls to OY_0. Since people are used to the standard of living at the OY_1 level of income, they will not reduce their consumption to E_0Y_0 level, but reduce it as little as possible by reducing their current saving. Thus they move backward along the C_{S1} curve to point C_1 and be at C_1Y_0 level of consumption. When the period of recovery starts, income rises to the previous peak level of OY_1. But consumption increases slowly from C_1 to E_1 along the C_{S1} curve because consumers will just restore their previous level of savings. If income continues to increase to OY_2 level, consumers will move upward along the

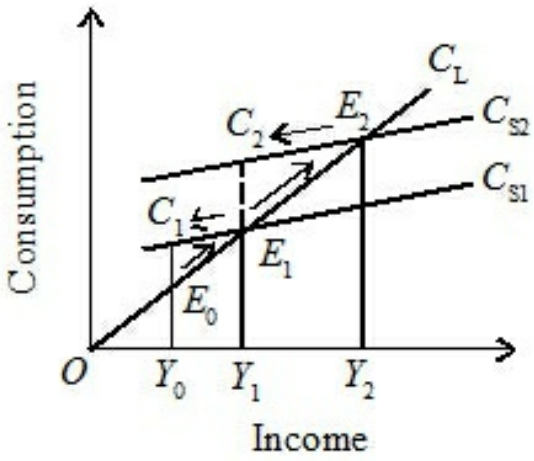


Fig 4

C_L curve from E_1 to E_2 on the new short-run consumption function C_{S2}. If another recession occurs at OY_2 level of income, consumption will decline along the C_{S2} consumption function toward C_2 point and income will be reduced to OY_1 level. But during recovery over the long-run, consumption will rise along the steeper C_L path till it reaches the short-run consumption function C_{S2}. This is because when income increases beyond its present level OY_1, the APC becomes constant over the long-run. The short-run consumption function shifts upward from Cs_1 to Cs_2 but consumers move along the C_L curve from E_1 to E_2. But when income falls, consumers move backward from E_2 to C_2 on the Cs_2 curve. These upward and downward movements from C_1 and C_2 points along the C_L curve give the appearance of a ratchet. This is the *rachet effect*. The short-run consumption function ratchets upward when income increases in the long run but it *does not shift down* to the earlier level when income declines. Thus the ratchet effect will develop *whenever* there is a cyclical decline or recovery in income.

Its Criticisms

Although the Duesenberry theory reconciles the apparent contradictions between budget studies and short-term and long-term time series studies, yet it is not without its deficiencies.

1. No Proportional Increase in Consumption. The relative income hypothesis assumes a proportional increase in income and consumption. But increases in income along the full employment level do not always lead to proportional increases in the consumption.

2. No Direct Relation between Consumption and Income. This hypothesis assumes the relation between consumption and income to be direct. But this has not been borne out by experience. Recessions do not always lead to decline in consumption, as was the case during the recessions of 1948-49 and 1974-75.

3. Distribution of Income not Unchanged. This theory is based on the assumption that the distribution of income remains almost unchanged with

the change in the aggregate level of income. If with increases in income, a redistribution occurs towards greater equality, the APC of all persons belonging to the relatively poor and relatively rich families will tend to be reduced. Thus the consumption function will not shift upward from C_{S1} to C_{S2} when income increases.

4. Reversible Consumer Behaviour. According to Micheal Evants, "The consumer behaviour is slowly reversible over time, instead of being truly ireversible. Then previous peak income would have less effect on current consumption, the greater the elapsed time from the last peak."[6] Even if we know how a consumer spent his previous peak income, it is not possible to know how he would spend it now.

5. Neglects Other Factors. This hypothesis is based on the assumption that changes in consumer's expenditure are related to his previous peak income. The theory is weak in that it neglects other factors that influence consumer spending such as asset holdings, urbanisation, changes in age-composition, the appearance of new consumer goods, etc.

6. Consumer Preferences do not Depend on Others. Another unrealistic assumption of the theory is that consumer preferences are interdependent whereby a consumer's expenditure is related to the consumption patterns of his rich neighbour. But this may not always be true. George Katona's[7] empirical study has revealed that expectations and attitudes play an important role in consumer spending. According to him, income expectations based on levels of aspirations and the attitudes toward asset holdings affect consumer spending behaviour more than the demonstration effect.

7. Reverse Lightning Bolt Effect. Smith and Jackson have criticised Duesenbery's empirical evidence that the recovery in income after recession is not caused by ratchet effect. Rather, the consumption experience of consumer is similar to the reverse lightning bolt effect. That is why the consumer gradually increases his consumption due to his inconsistent habit stability with the increase in his income after recession. This is shown is Fig.5 where the levels of consumption with the increments in income have been shown by arrows as reverse lightning bolt

takes place.

6. Micheal K. Evans, *Macroeconomic Activity*, 1969

7. George Katona, *Psychological Analysis of Economic Behaviour*, 1963

8. M. Friedman, *A Theory of Consumption Function*, 1957

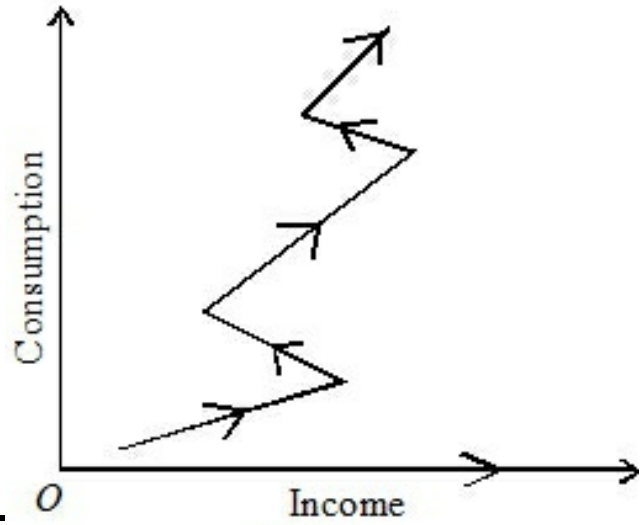


Fig 5

The Permanent Income Hypothesis

Another solution to the apparent contradiction between the proportional long-run and non-proportional short-run consumption function is Friedman's[8] permanent income hypothesis. Friedman rejects the use of "current income" as the determinant of consumption expenditure and instead divides both consumption and income into "permanent" and "transitory" components, so that

$$Ym \text{ or } Y = Y_P + Y_t \qquad ...(1)$$

and

$$C = C_P + C_t \qquad ...(2)$$

where *p* refers to permanent, *t* refers to transitory, *Y* to income and *C* to consumption.

Permanent income is defined as "the amount a consumer unit could consume (or believes that it could) while maintaining its wealth intact." It is the main income of a family unit which in turn depends on its time-horizon and farsightedness. "It includes non-human wealth that it owns, the personal attributes of earners in the unit...the attributes of the economic activity of the earners such as the occupation followed, the location of economic activity, and so on."

Y being the consumer's measured income or current income, it can be larger or smaller than his permanent income in any period. Such differences between measured and permanent income are due to the transitory component of income (Y_t). Transitory income may rise or fall with windfall gains or losses and cyclical variations. If the transitory income is *positive* due to a windfall gain, the measured income will rise above the permanent income. If the transitory income is *negative* due to theft, the measured income falls below the permanent income. The transitory income can also be *zero* in which case measured income equals permanent income.

Permanent consumption is defined as "the value of the services that it is planned to consume during the period in question." Measured consumption is also divided into permanent consumption (C_p) and transitory consumption (C_t). Measured consumption (or current consumption) may deviate from or equal permanent consumption depending on whether the transitory consumption is positive, negative or zero, Permanent consumption (C_p) is a multiple (k) of permanent income, Y_p .

$$C_P = kY_P$$

and $$k = f(r, w, u)$$

Therefore, $$C_P = k(r, w, u)\, Y_P \qquad ...(3)$$

where k is a function of the rate of interest (r), the ratio of property and non-property income to total wealth or national wealth (w), and the consumer's propensity to consume (u). This equation tells that over the long period consumption increases in proportion to the change in Y_P. This is attributable to a constant k ($=C_p / Y_p$) which is independent of the size of income. Thus k is the permanent and average propensity to consume and *APC = MPC*.

Friedman analyses the offsetting forces which lead to this result. To take the rate of interest (r), there has been a secular decline in it since the

1920s. This tends to raise the value of *k*. But there has been a long-run decline in the ratio of property and non-property income to national wealth (*w*) which tends to reduce the value of *k*. The propensity to consume has been influenced by three factors. *First*, there has been a sharp decline in the farm population which has tended to increase consumption with urbanisation. This has led to increase of *k*. *Second*, there has been a sharp decline in the size of families. It has led to increase in saving and reduction in consumption thereby reducing the value of *k*. *Third*, larger provision by the state for social security. This has reduced the need for keeping more in savings. It has increased the tendency to consume more resulting in the rise in the value of *k*. The overall effect of these off-setting forces is to raise consumption in proportion to the change in the permanent income component.

Therefore, there is a proportional relation between permanent income and consumption,

$$C_p = kY_p \qquad ...(4)$$

where *k* is the coefficient of proportionality in which *APC* and *MPC* are endogenous and it depends upon the above mentioned factors. In other words, it is that proportion of fixed income which is consumed. Now take permanent income which is based on time series. Friedman believes that permanent income depends partly on current income and partly on previous period's income. This can be measured as

$$Y_{pt} = aY_t + (1–a)\ Y_{t–1} \qquad ...(5)$$

where Y_{pt} = permanent income in the current period, Y_t = current income in the current period, $Y_{t\text{-}1}$ = previous period's income, a = ratio of change in income between current period (t) and previous period (t–1).

This equation tells that permanent income is the sum of current period's income (Y_t) and previous periods income ($Y_{t–1}$) and the ratio of income change between the two (a). If the current income increases at once, there will be small increase in permanent income. For the permanent income to

increase, income will have to be raised continuously for many years. Then only people will think that it has increased.

By integrating equations (4) and (5), short-run and long-run consumption function can be explained as

$$C_t = kY_{pt} = kaY_t + k\,(1–a)Y_{t\text{-}1} \qquad ...(6)$$

where C_t = current period consumption, ka = short-run *MPC*, k = long-run *MPC* and $k\,(1–a)\,Y_{t\text{-}1}$ is the intercept of short-run consumption function.

According to Friedman, k and ka are different from one another and $k > ka$. Further, $k = 1$ and $ka = 0$

Equation (6) tells that consumption depends both on previous income and current income. Previous income is important for consumption because it helps in forecasting the future income of people.

Its Assumptions

Given these, Friedman gives a series of assumptions concerning the relationships between permanent and transitory components of income and consumption.

1. There is no correlation between transitory income and permanent income.

2. There is no correlation between permanent and transitory consumption.

3. There is no correlation between transitory consumption and transitory income.

4. Only differences in permanent income affect consumption systematically.

Explanation of the Theory

These assumptions give the explanation of the cross-section results of

Friedman's theory that the short-run consumption function is linear and non-proportional, *ie*. *APC* > *MPC* and the long-run consumption function is linear and proportional, *ie*. *APC* = *MPC*.

Figure 6 explains the permanent income hypothesis of Friedman where C_L is the long-run consumption function which represents the long-run proportional relationship between consumption and income of an individual where *APC* = *MPC*. C_S is the non-proportional short-run consumption function where measured income includes both permanent and transitory components. At *OY* income level where C_S and C_L curves coincide at point *E*, permanent income and measured income are identical and so are permanent and measured consumption as shown by *YE*. At point *E*, the transitory factors are non-existent. If the consumer's income increases to OY_1, he will increase his consumption consistent with the rise in his income. For this, he will move along the C_S curve to E_2 where his measured income in the short-run is OY_1 and measured consumption is Y_1E_2. The reason for this movement from *E* to E_2 is that during the short-run the consumer does not expect the rise in income to be permanent, so *APC* falls as income increases. But if the OY_1 income level becomes permanent, the consumer will also increase his consumption permanently. Now his short-run consumption function will shift upward from C_S to C_{S1} and intersect the long-run consumption function C_L at point E_1. Thus the consumer will consume Y_1E_1 at OY_1 income level. Since he knows that the increase in his income OY_1 is permanent, he will adjust his consumption Y_1E_1 accordingly on the long-run consumption function C_L at E_1 where *APC* = *MPC*

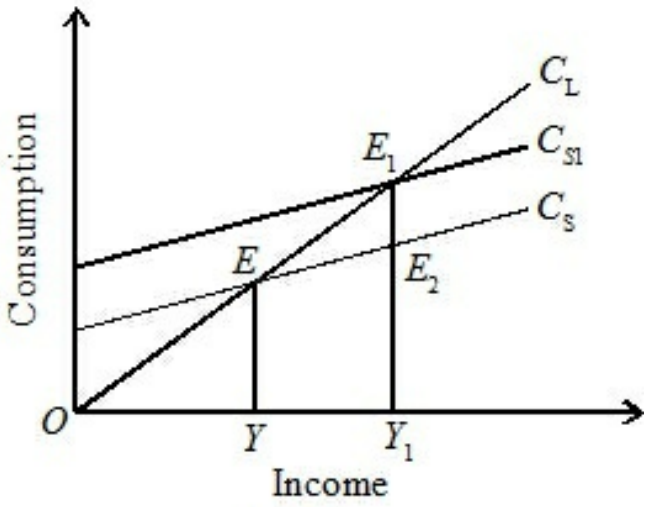


Fig 6

Its Criticisms

This theory has been criticised on the following counts:

1. Correlation between Temporary Income and Consumption. Friedman's assumption that there is no correlation between transitory components of consumption and income is unrealistic. This assumption implies that with the increase or decrease in the measured income of the household, there is neither any increase or decrease in his consumption, because he either saves or dissaves accordingly. But this is contrary to actual consumer behaviour. A person who has a windfall gain does not deposit the entire amount in his bank account but enjoys the whole or part of it on his current consumption. Similarly, a person who has lost his purse would definitely cut or postpone his present consumption rather than rush to the bank to withdraw the same amount of money to meet his requirements.

2. APC of all Income Groups not Equal. Friedman's hypothesis states that the APC of all families, whether rich or poor, is the same in the long-run. But this is against the ordinary observed behaviour of households. It is an established fact that low-income families do not have the capacities to save the same fraction of their incomes as the high income families. This is not only due to their meagre incomes but their tendency to prefer present consumption to future consumption in order to meet their unfulfilled wants. Therefore, the consumption of low-income families is higher relative to their incomes while the saving of high-income families is higher relative to their incomes. Even in the case of persons at the same level of permanent income, the level of saving differs and so does consumption.

3. Use of Various terms Confusing. Friedman's use of the terms "permanent", "transitory", and "measured" have tended to confuse the theory. The concept of measured income improperly mixes together permanent and transitory income on the one hand, and permanent and transitory consumption on the other.

4. No Distinction between Human and Non-human Wealth. Another weakness of the permanent income hypothesis is that Friedman does not make any distinction between human and non-human wealth and includes income from both in a single term in the empirical analysis of his theory.

Conclusion. Despite these weaknesses, "it can be fairly said", according to Micheal Evans, "that the evidence supports this theory *and that* Friedman's formulation has reshaped and redirected much of the research on the consumption function."[9]

THE LIFE CYCLE HYPOTHESIS

Ando and Modigliani[10] have formulated a consumption function which is known as the Life Cycle Hypothesis*. According to this hypothesis, consumption is a function of lifetime expected income of the consumer. The consumption of the individual consumer depends on the resources available to him, the rate of return on capital, the spending plan, and the age at which the plan is made. The present value of his resources includes income from assets or wealth or property and from *current and expected* labour income. Thus his total resources consist of his income and wealth.

---

9. M.K.Evans, *op.cit*., p.34 Italics added.

10. A. Ando and F. Modigliani, "The Life Cycle Hypothesis of Saving : Aggregate Implications and Tests," *A.E.R*., March 1963.

* This is also known as the Wealth Theory of Consumption.

Its Assumptions

The life cycle hypothesis is based on the following assumptions:

1. There is no change in the price level during the life of the consumer.

2. The rate of interest paid on assets is zero.

3. The consumer does not inherit any assets and his net assets are the result of his own savings.

4. His current savings result in future consumption.

5. He intends to consume his total lifetime earnings plus current assets.

6. He does not plan any bequests.

7. There is certainty about his present and future flow of income.

8. The consumer has a definite conscious vision of life expectancy.

9. He is aware of the future emergencies, opportunities and social pressures which will impinge upon his consumption spending.

10. The consumer is rational.

Its Explanation

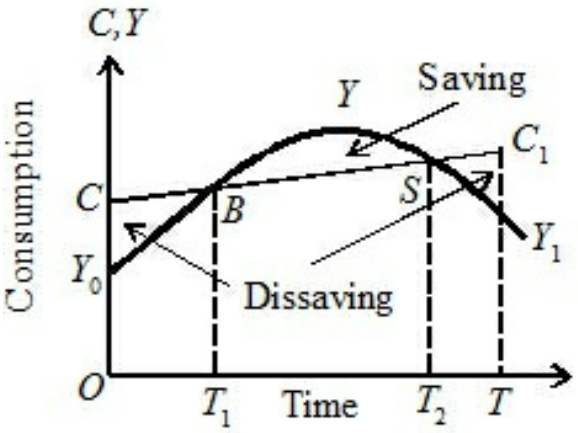


Fig 7

Given these assumptions, the aim of the consumer is to maximise his utility over his lifetime which will, in turn, depend on the total resources available to him during his lifetime. Given the life-span of an individual, his consumption is proportional to these resources. But the proportion of resources that the consumer plans to spend will depend on whether the spending plan is formulated during the early or later year of his life. As a rule, an individual's average income is relatively low at the beginning of his life and also at the end of his life. This is because in the early years of his life, he has little assets (wealth) and during the late years, his labour-income is low. It is, however, in the middle of his life that his income, both from assets and labour, is high. As a result, the consumption level of the individual throughout his life is somewhat constant or slightly increasing, shown as the CC_1 curve in Fig. 7, the Y_0YY_1 curve shows the individual consumer's income stream during his lifetime T. During the early period of his life represented by T_1 in the figure., he borrows or dissaves CY_0B amount of money to keep his consumption level CB which is almost constant. In the middle years of his life represented by T_1T_2, he saves BSY amount to repay his debt and for the future. In the last years of his life represented by T_2T, he dissaves SC_1Y_1 amount.

According to this theory*, consumption is a function of lifetime expected income of the consumer which depends on his resources. In some resources, his current income (Y_t); present value of his future expected labour income (Y^e_{Lt}) and present value of assets (A_t) are included.

---

* This portion can be left out without loss in continuity by ordinary student.

The consumption function can be expressed as :

$$C_t = f(V_t) \qquad ...(1)$$

where V_t = total resources at time *t*.

and $V_t = f(Y_t + Y^e_{Lt} + A_t)$ ...(2)

By substituting equation (2) in (1) and making (2) linear and weighted average of different income groups, the aggregate consumption function is

$$C_t = \alpha_1 Y_t + \alpha_2 Y^e_L + \alpha_3 A_t \qquad ...(3)$$

where α_1 = *MPC* of current income, α_2 = *MPC* of expected labour income; and α_3 = *MPC* of assets or wealth.

Now *APC* is

$$\frac{C_t}{Y_t}\alpha_1 + \alpha_2 \frac{Y^e_L}{Y_t} + \alpha_3 \frac{A_t}{Y_t}$$

APC is constant in the long-run because a portion of labour income in current income and the ratio of total assets to current income are constant when the economy grows.

On the basis of the life cycle hypothesis, Ando and Modigliani made a number of studies in order to formulate the short-run and long-run consumption functions. A cross-section study revealed that more persons

in the low-income groups were at *low* income level because they were at the end period of their lives. Thus their *APC* was high. On the other hand, more than average persons belonging to the high-income groups were at *high* income levels because they were in the middle years of their lives. Thus their *APC* was relatively low. On the whole, the *APC* was falling as income rose thereby showing *APC*> *MPC*. The observed data for the U.S. revealed the *APC* to be constant at 0.7 over the long-run.

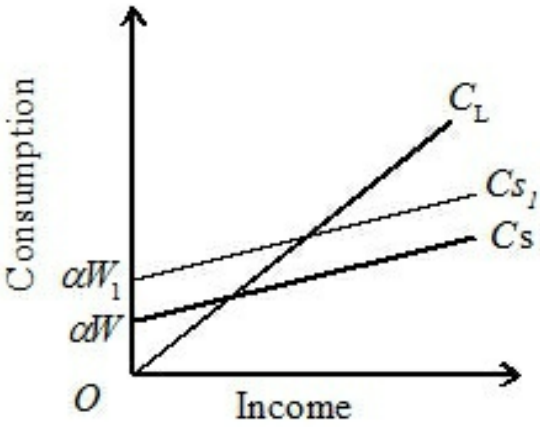


Fig 8

The Ando-Modigliani short-run consumption function is shown by the C_S curve in Fig. 8. At any given point of time, the C_S curve can be considered as a constant and during short-run income fluctuation, when wealth remains fairly constant, it looks like the Keynesian consumption function. But its intercept will change as a result of accumulation of wealth (assets) through savings As wealth increases overtime, the non-proportional short-run consumption function C_S shifts upward to C_{S1} to trace out the long-run proportional consumption function. The long-run consumption function is C_L, showing a constant *APC* as income grows along the trend. It is a straight line passing through the origin. The *APC* is constant over time because the share of labour income in total income and the ratio of wealth (assets) to total income are constant as the economy grows along the trend.

Its Implications

1. The life cycle hypothesis solves the *consumption puzzle*. According to this hypothesis, the short-run consumption function would be non-proportional as in the short-run time series estimates. Its intercept (αW in Fig. 8) measures the effect of wealth and the life cycle consumption function looks like the Keynesian consumption function as C_S in the figure But it holds only in the short run when wealth is constant. As

wealth grows (αW_1), this consumption function shifts upward as C_{S1}. The shifting of the C_S to C_{S1} traces out the long-run consumption function, C_L. This is consistent with the evidence from long-run time series data that the long-run consumption function is proportional. The slope of the C_L curve shows that the average propensity to consume does not fall as income increases. In this way, Audo-Modigliani solved the consumption puzzle.

2. The life cycle hypothesis reveals that *savings change over the life time* of a consumer. If a consumer starts his life in adulthood with no wealth, he will save and accumulate wealth during his working years. But during retirement, he will dissave and run down his wealth. Thus the life cycle hypothesis implies that the consumer wants smooth and uninterrupted consumption over his lifetime. During working years, he saves and when retires, he dissaves.

3. The life cycle hypothesis also implies that a *high-income family consumes* a *smaller proportion* of *his income* than a low-income family. In its peak earning years, (shown as portion BSY in Fig. 7), its income is more than its consumption and its APC is the lowest. But in the case of a low-income family and a retiree family, the APC is high.

Its Criticisms

The life cycle hypothesis is not free from certain criticisms.

1. Plan for Lifetime Consumption Unrealistic. The contention of Audo and Modigliani that a consumer plans his consumption over his lifetime is unrealistic because a consumer concentrates more on the present rather than on the future which is uncertain.

2. Consumption not directly related to Assets. The life cycle hypothesis pre-supposes that consumption is directly related to the assets of an individual. As assets increase, his consumption increases and vice versa. This is also unwarranted because an individual may reduce his consumption to have larger assets.

3. Consumption depends on Attitude. Consumption depends upon one's

attitude towards life. Given the same income and assets, one person may consume more than the other.

4. Consumer not Rational and Knowledgeable. This hypothesis assumes that the consumer is rational and has full knowledge about his income and future litetime. This is unrealistic because no consumer is fully rational and knowledgeable.

5. Estimation of Variables not Possible. This theory depends on many variables such as current income, value of assets, future expected labour income, etc., the estimation of so many variables is very difficult and not possible.

6. Liquidity Constraints. This hypothesis fails to recognise the existence of liquidity constraints for a consumer. Even if he possesses a definite and conscious vision of future income, he may have little opportunity for borowing in the capital market on the basis of expected future income. As a result, consumption may response more to changes in current income than predicted on the basis of the life cycle hypothesis.

Conclusion. Despite these, the life cycle hypothesis is superior to the other hypotheses on consumption function because it includes not only wealth as a variable in the consumption function but also explains why *APC* > *MPC* in the short-run and *APC* is constant in the long-run.

SELECTED READINGS

Edward Shapiro, *Macroeconomics Analysis*, 4/e, 1978, Ch. 8.

J.H. Makin, *Macroeconomics*, 1975, Ch. 6.

P. Wonnacott, *Macroeconomics*, 1974, Ch. 13.

Norman K. Evans, *Macroeconomic Activity*, 1969, Ch. 2.

EXERCISES

1. What differences are generally noticed between the short-run and long-run behaviour of the consumption function ? How do you explain these differences?
2. Explain Keynes' Absolute Income Hypothesis. Discuss the evidences that support and that do not support his hypothesis.
3. Explain fully the consumption puzzle.
4. What is the Relative Income Hypothesis for explaining the aggregate consumption behaviour ? Is it an improvement on the Absolue Income Hypothesis ?
5. Discuss critically Friedman's Permanent Income Hypothesis of Consumption.
6. Examine critically the Life Cycle Consumption Hypothesis.
7. Discuss the Drift Theory of Consumption.

CHAPTER 11

THE INVESTMENT FUNCTION

MEANING OF CAPITAL AND INVESTMENT

In ordinary parlance, investment means to buy shares, stocks, bonds and securities which are already existing in stock market. But this is not real investment because it is simply a transfer of existing assets. Hence this is called financial investment which does not affect aggregate spending. In Keynesian terminology, investment refers to real investment which adds to capital equipment. It leads to increase in the levels of income and production by increasing the production and purchase of capital goods. Investment thus includes new plant and equipment, construction of public works like dams, roads, buildings, etc., net foreign investment, inventories and stocks and shares of new companies. In the words of Joan Robinson, "By investment is meant an addition to capital, such as occurs when a new house is built or a new factory is built. Investment means making an addition to the stock of goods in existence."

Capital, on the other hand, refers to real assets like factories, plants, equipment, and inventories of finished and semi-finished goods. It is any previously produced input that can be used in the production process to produce other goods. The amount of capital available in an economy is the stock of capital. Thus capital is a *stock* concept.

To be more precise, investment is the production or acquisition of real capital assets during any period of time. To illustrate, suppose the capital assets of a firm on 31 March 2004 are Rs 100 crores and it invests at the rate of Rs 10 crores during the year 2004-05. At the end of the next year

(31 March 2005), its total capital will be Rs 110 crores. Symbolically, let I be investment and K be capital in year t, then $I_t = K_t - K_{t-1}$.

Capital and investment are related to each other through net investment. Gross investment is the total amount spent on new capital assets in a year. But some capital stock wears out every year and is used up for depreciation and obsolescence. Net investment is gross investment *minus* depreciation and obsolescence charges for replacement investment. This is the net addition to the existing capital stock of the economy. If gross investment equals depreciation, net investment is zero and there is no addition to the economy's capital stock. If gross investment is less than depreciation, there is disinvestment in the economy and the capital stock decreases. Thus for an increase in the real capital stock of the economy, gross investment must exceed depreciation, i.e., there should be net investment.

TYPES OF INVESTMENT

Induced Investment. Real investment may be induced. Induced investment is profit or income motivated. Factors like prices, wages and interest change which affect profits and influence induced investment. Similarly, demand also influences it. When income increases, consumption demand also increases and to meet this, investment increases. In the ultimate analysis, induced investment is a function of income i.e., $I = f(Y)$. It is income elastic. It increases or decreases with the rise or fall in income, as shown in Figure 1.

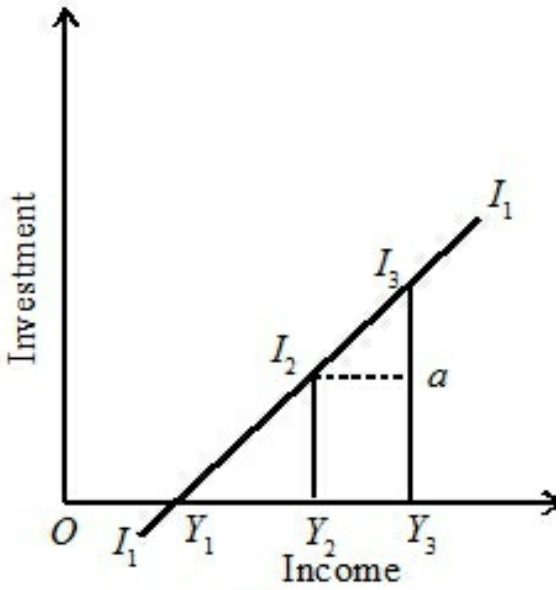


Fig. 1

I_1I_1 is the investment curve which shows induced investment at various

levels of income. Induced investment is zero at OY_1 income. When income rises to OY_3, induced investment is I_3Y_3. A fall in income to OY_2 also reduces induced investment to I_2Y_2.

Induced investment may be further divided into (*i*) the average propensity to invest, and (*ii*) the marginal propensity to invest:

(i) The average propensity to invest is the ratio of investment to income, I/Y. If the income is Rs. 40 crores and investment is Rs. 4 crores, I/Y = 4/40 = 0.1. In terms of the above figure, the average propensity to invest at OY_3 income level is I_3Y_3 / OY_3.

(ii) The marginal propensity to invest is the ratio of change in investment to the change in income, i.e., $\Delta I/\Delta Y$. If the change in investment, ΔI=Rs 2 crores and the change in income, ΔY = Rs 10 crores, then $\Delta I/\Delta Y$ = 2/10=0.2 In Figure 1, $\Delta I/\Delta Y = I_3a/Y_2Y_3$.

Autonomous Investment. Autonomous investment is independent of the level of income and is thus income inelastic. It is influenced by exogenous factors like innovations, inventions, growth of population and labour force, researches, social and legal institutions, weather changes, war, revolution, etc. But it is not influenced by changes in demand. Rather, it influences demand. Investment in economic and social overheads whether made by the government or the private enterprise is autonomous. Such investment includes expenditures on buildings, dams, roads, canals, schools, hospitals, etc. Since investment on these projects is generally associated with public policy, autonomous investment is regarded as public investment. In the long-run, private investment of all types may be autonomous because it is influenced by exogenous factors. Diagrammatically, autonomous investment is shown as a curve parallel to the horizontal axis as I_1I' curve in Figure 2. It indicates

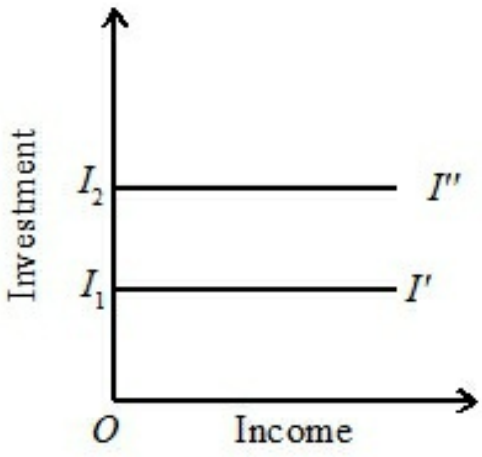


Fig. 2

that at all levels of income, the amount of investment OI_1 remains constant. The upward shift of the curve to I_2I'' indicates an increased steady flow of investment at a constant rate OI_2 at various levels of income. However, for purposes of income determination, the autonomous investment curve is superimposed on the C curve in a 45° line diagram.

THE PRESENT VALUE (PV) CRITERION OF INVESTMENT

The present value criterion is considered to be the best method for evaluating capital investment proposals. Profitability of an investment project is evaluated by this method. It is also called as the net present value (NPV) criterion. It is calculated by using an appropriate rate of interest which is the capital cost of a firm. This is the minimum rate of expected return likely to be earned by the firm on investment proposals. To find out the present value of cash flows expected in future periods, all the cash outflows and cash inflows, are discounted at the above rate. Net present value is the difference between total present value of cash outflows and total present value of cash inflows occurring in periods over the entire life of the project. When the net present value is positive, the investment proposal is profitable and worth selecting. But if it is negative, the investment proposal is non-profitable and rejectable. To calculate net present value index of different investment proposals, the following method can be used:

$$\text{NPV} = \frac{\text{Total Present Value of All Cash Flows}}{\text{Initial Investment}}$$

NPV method considers the time value of money. It compares time value of cash flows.

NPV = Present value of Gross Earnings – Net Cash Investment

NPV can be found out from the following formula :

$$\text{NPV} = \frac{A_1}{(1+r)^1} + \frac{A_2}{(1+r)^2} + \frac{A_3}{(1+r)^3} \ldots\ldots \frac{A_n}{(1+r)^n} - C$$

Where A_1, A_2, A_3, etc. are the cash inflows at the end of first, second and third year respectively

n = Expected life of investment proposals.

r = Rate of discount which is equal to the cost of capital.

C = Present value of costs.

Thus, NPV = Sum of Discounted Gross Earnings – Sum of Discounted Value of Cost

For example, if the initial investment cost of a project is Rs. 100 crores, cash inflow in the coming years is Rs. 125 crores and the market rate of interest is 10% p.a., NPV will be as follows :

$$\text{NPV} = \frac{125}{(1+0.10)^1} - 100$$

$$= 125 \times \frac{10}{11} - 100$$

$$= 113.64 - 100 = 13.64$$

Accept or Reject Criterion

The decision criterion relating to NPV may be as under:

(a) If NPV > 0, the project is profitable.

(b) If NPV < 0, the project will not be profitable.

(c) If NPV = 0, the project may or may not be started.

If the decision is to be taken between two investment projects, the project with high positive NPV would be selected rather than the other.

Merits

NPV method has the following merits:

(i) This method considers time value of money.

(ii) It considers the cash flows of the project in different time periods.

(iii) It is more scientific than traditional methods.

(iv) NPVs of different projects can be added to arrive at the cumulative NPV for a business.

(v) Intermediate cash flows are reinvested at the discount rate.

(vi) The calculation of NPV allows expected change in discount rate.

(vii) The discount rate (r) used for discounting future cash flows is, in fact, the minimum necessary rate of return which consists of both net of return and premium required to set off the risk.

(viii) The discount rate applied in the NPV method is the capital cost of the firm.

(ix) The use of NPV is the most profitable in view of the maximum profit.

Demerits

There are also some demerits of this method:

(i) It is difficult to calculate the profit cost with this method.

(ii) It is difficult to work out especially the cost of equity capital by this method.

(iii) It is not applicable without the knowledge of cost of capital.

(iv) It favours long-run investment projects.

(v) When projects with different investments are compared, this method does not give correct result.

(vi) Its assumption that the intermediate cash flows are reinvested on the capital cost of the firms, is not always true.

(vii) This method gives different rankings in the case of complicated projects in comparison to other methods.

DETERMINANTS OF THE LEVEL OF INVESTMENT

The decision to invest in a new capital asset depends on whether the expected rate of return on the new investment is equal to or greater or less than the rate of interest to be paid on the funds needed to purchase this asset. It is only when the expected rate of return is higher than the interest rate that investment will be made in acquiring new capital assets.

In reality, there are three factors that are taken into consideration while making any investment decision. They are (a) the cost of the capital asset, (b) the expected rate of return from it during its lifetime, and (c) the market rate of interest. Keynes sums up these factors in his concept of the marginal efficiency of capital (*MEC*).

Marginal Efficiency of Capital (MEC)

The marginal efficiency of capital is the highest rate of return expected from an additional unit of a capital asset over its cost. In the words of Kurihara, “It is the ratio between the prospective yield of additional capital goods and their supply price.” The prospective yield is the aggregate net return from an asset during its life time, while the supply price is the cost of producing this asset. If the supply price of a capital asset is Rs. 20,000 and its annual yield is Rs. 2,000, the marginal efficiency of this asset is $\frac{2000}{20000} \times \frac{100}{1}$ = 10 per cent. Thus the marginal

efficiency of capital (MEC) is the *percentage of profit expected from a given investment on a capital asset.*

Keynes relates the prospective yield of a capital asset to its supply price and defines the *MEC* as "equal to the rate of discount which would make the present value of the series of annuities given by the returns expected from the capital assets during its life just equal to its supply price."[1]

Symbolically, this can be expressed as:

$$S_p = \frac{R_1}{(1+i)^1} + \frac{R_2}{(1+i)^2} + \ldots + \frac{R_n}{(1+i)^n} \qquad \ldots(1)$$

Where *Sp* is the supply price or the cost of the capital asset; R_1, R_2... and R_n are the prospective yields or the series of expected annual returns from the capital asset in the years, 1,2... and *n*; and *i* is the rate of discount which makes the capital asset exactly equal to the present value of the expected yield from it. This *i* is the *MEC* or the rate of discount which equates the two sides of the equation. If the supply price of a new capital asset is Rs 1,000 and its life is two years, it is expected to yield Rs 550 in the first year and Rs 605 in the second year. Its *MEC* is 10 per cent which equates the supply price to the expected yields of this capital asset.

Thus

$$(Sp)\ \text{Rs } 1000 = \frac{550}{(110)} + \frac{605}{(110)^2} = \text{Rs. } 500 + 500$$

In equation (1), the term $\frac{R_1}{(1+i)}$ is the present value (*PV*) of the capital asset. The present value is "the value of payments to be received in the future." It depends on the rate of interest at which it is discounted.

Suppose we expect to receive Rs 100 from a machine in a year's time and the rate of interest is 5 per cent. The present value of this machine is

---

1. J.M. Keynes, *op. cit.*, p. 135.

$$\frac{R_1}{(1+i)} = \frac{100}{(1.05)} = \text{Rs } 95.24.$$

If we expect Rs 100 from the machine after two years then its present value is $\frac{100}{(1.05)^2}$ = Rs 90.70. The present value of a capital asset is inversely related to the rate of interest. The lower the rate of interest, the higher is the present value, and vice versa. For instance, if the rate of interest is 5 per cent, *PV* of an asset of Rs 100 for one year will be Rs 95.24; at 7 per cent interest rate, it will be Rs 93.45; and at 10 per cent interest rate, it will be Rs 90.91.

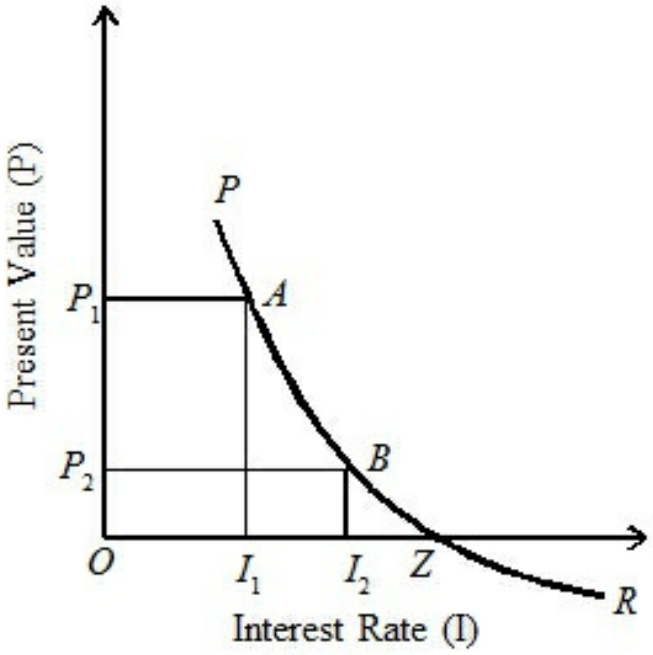


Fig. 3

The relation between the present value and the rate of interest is shown in Figure 3, where the rate of interest is taken on the horizontal axis while the present value of the project on the vertical axis. The curve *PR* shows the inverse relation between present value and rate of interest. If the current rate of interest is i_1,, the present value of the project is P_1. On the other hand, a higher rate of interest (I_2) will lead to a lower present value (P_2) and when the present value curve (*PR*) cuts the horizontal axis at point (*Z*), the net present value becomes zero.

As a matter of fact, the *MEC* is the expected rate of return over cost of a new capital asset. In order to find out whether it is worthwhile to purchase a capital asset, it is essential to compare the present value of the capital asset with its cost or supply price. If the present value of a capital asset exceeds its cost of buying, it pays to buy it. On the contrary, if its present value is less than its cost, it is not worthwhile investing in this capital asset.

The same results can be had by comparing the *MEC* with the market rate of interest. If the *MEC* of a capital asset is higher than the market rate of interest at which it is borrowed, it pays to purchase the capital asset, and vice versa. If the market interest rate equals the *MEC* of the capital asset, the firm is said to possess the *optimum* capital stock. If the *MEC* is higher than the rate of interest, there will be a tendency to borrow funds in order to invest in new capital assets. If the *MEC* is lower than the rate of interest, no firm will borrow to invest in capital assets. Thus the equilibrium condition for a firm to hold the optimum capital stock is where the *MEC* equals the interest rate. Any disequilibrium between the *MEC* and the rate of interest can be removed by changing the capital stock, and hence the *MEC* or by changing the rate of interest or both. Since the stock of capital changes slowly, therefore, changes in the rate of interest are more important for bringing equilibrium. The above arguments which have been applied to a firm are equally applicable to the economy.

Figure 4 shows the *MEC* curve of an economy. It has a negative slope (from left to right downward) which indicates that the higher the *MEC*, the smaller the capital stock. Or, as the capital stock increases, the *MEC* falls. This is because of the operation of the law of diminishing returns in production. As a result, the marginal physical productivity of capital and the marginal revenue fall. In the figure, when the capital stock is OK_1, the *MEC* is Or_1. As the capital increases from OK_1 to OK_2 the *MEC* falls from Or_1 to Or_2. The net addition to the capital stock K_1K_2 represents the net investment in the economy.

Further, to reach the optimum (desired) capital stock in the economy, the *MEC* must equal the rate of interest. If, as shown in the figure, the existing

capital stock is OK_1, the *MEC* is Or_1 and the rate of interest is at Or_2. Everyone in the economy will borrow funds and invest in capital assets. This is because *MEC* (Or_1) is higher than the rate of interest (at Or_2). This will continue till the *MEC* (Or_1) comes down to the level of the interest rate (at Or_2). When the *MEC* equals the rate of interest, the economy reaches the level of optimum capital stock. The fall in the *MEC* is due to the increase in the actual capital stock from OK_1 to the optimum (desired) capital stock OK_2. The increase in the firm's capital stock by K_1K_2 is the net investment of the firm. But it is the rate of interest which determines the size of the optimum capital stock in the economy. And it is the *MEC* which relates the amount of desired capital stock to the rate of interest. Thus the negative slope of the *MEC* curve indicates that as the rate of interest falls the optimum stock of capital increases.

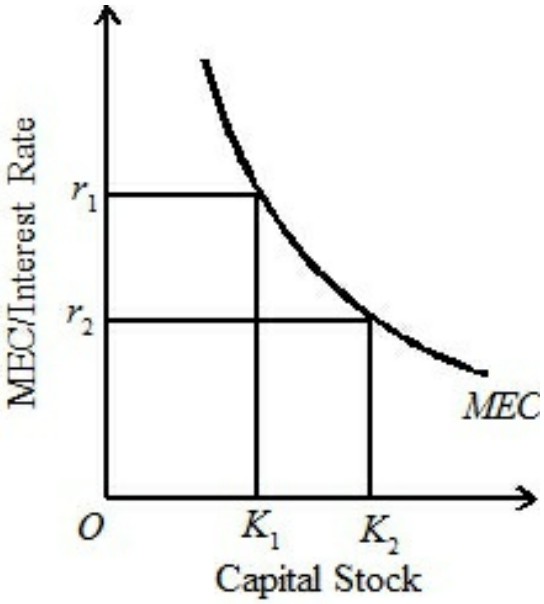


Fig. 4

Marginal Efficiency Of Investment (*MEI*)

The marginal efficiency of investment is the rate of return expected from a given investment on a capital asset after covering all its costs, except the rate of interest. Like the *MEC*, it is the rate which equates the supply price of a capital asset to its prospective yield. The

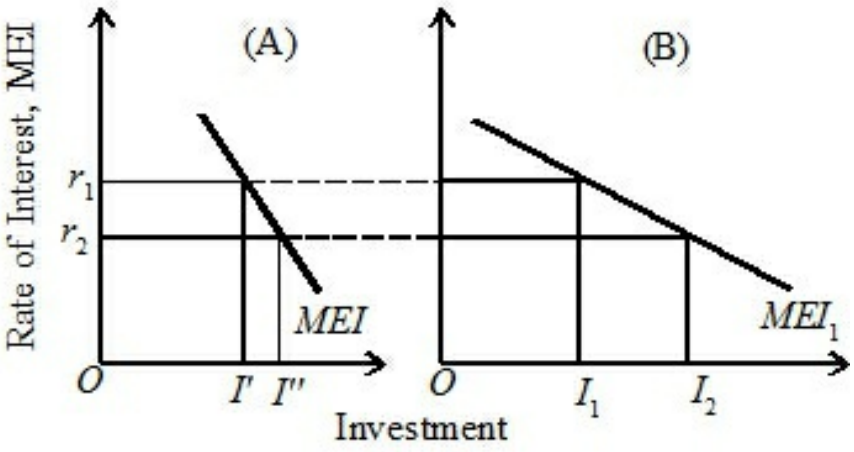


Fig. 5

investment on an asset will be made depending upon the interest rate involved in getting funds from the market. If the rate of interest is high, investment is at a low level. A low rate of interest leads to an increase in investment. Thus the *MEI* relates investment to the rate of interest. The *MEI* schedule shows the amount of investment demanded at various rates of interest. That is why, it is also called the investment demand schedule or curve which has a negative slope, as shown in Fig. 5(A). At Or_1 rate of interest, investment is *OI'*. As the rate of interest falls to Or_2, investment increases to *OI''*.

To what extent the fall in the interest rate will increase investment depends upon the elasticity of the investment demand curve or the *MEI* curve. The less elastic is the *MEI* curve, the lower is the increase in investment as a result of fall in the rate of interest, and vice versa.

In Figure 5. the vertical axis measures the interest rate and the *MEI* and the horizontal axis measures the amount of investment. The *MEI* and MEI_1 are the investment demand curves. The *MEI* curve in Panel (A) is less elastic to investment which increases by *I'I''*. This is less than the increase in investment I_1I_2 shown in Panel (B) where the MEI_1 curve is elastic. Thus given the shape and position of the *MEI* curve, a fall in the interest rate will increase the volume of investment.

On the other hand, given the rate of interest, the higher the *MEI*, the larger shall be the volume of investment. The higher marginal efficiency of investment implies that the *MEI* curve shifts to the right. When the existing capital assets wear out, they are replaced by new ones and the level of investment increases. But the amount of induced investment depends on the existing level of total purchasing. So more induced investment occurs when the total purchasing is higher. The higher total

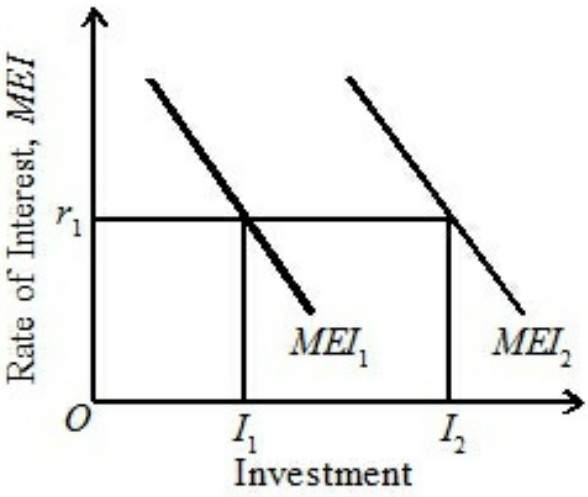


Fig. 6

purchasing tends to shift the *MEI* to the right indicating that more inducement to investment takes place at a given level of interest rate. This is explained in Figure 6, where MEI_1 and MEI_2 curves indicate two different levels of total purchasing in the economy. Let us suppose that the MEI_1 curve indicates that at Rs 200 crores of total purchasing, OI_1 (Rs 20 crores) investment occurs at Or_1 interest rate. If total purchasing rises to Rs 500 crores, the MEI_1 curve shifts to the right as MEI_2 and the level of induced investment increases to OI_2 (Rs 50 crores) at the same interest rate Or_1.

Relation between the MEC (Capital Stock) and the MEI (Investment)

Professor Lerner pointed out as early as in 1946 that Keynes erred not only descriptively but also analytically by failure to distinguish between the marginal efficiency of capital (*MEC*) and the marginal efficiency of investment (*MEI*). Following Lerner[2], Gardner Ackley[3] and some other economists have clearly defined and distinguished between the two concepts.

The *MEC is* based on a given supply price for capital, and the *MEI* on induced changes in this price. The *MEC* shows the rate of return on all successive units of capital without regard to the existing stock of capital. On the other hand, the *MEI* shows the rate of return on only units of capital over and above the existing stock of capital. In the *MEC,* the capital stock is taken on the horizontal axis of a diagram, while in the *MEI* the amount of investment is taken horizontally on the *X*-axis. The former is a ‘stock’ concept, and the latter is a ‘flow’ concept.

The *MEC* determines the optimum capital stock in an economy at each level of interest rate. The *MEI* determines the net investment of the economy at each interest rate, given the capital stock. The net investment is the addition to the existing capital stock whereby the actual capital stock increases. Investment will, therefore, continue to be made in the economy till the optimum capital stock is reached. The amount of investment to be made to attain the optimum capital stock in the economy will depend upon the law of production under which the capital goods

industry is operating. Under the law of constant costs, the supply curve of capital will be perfectly elastic, and the rate of investment per time period is determined as shown in figure 7.

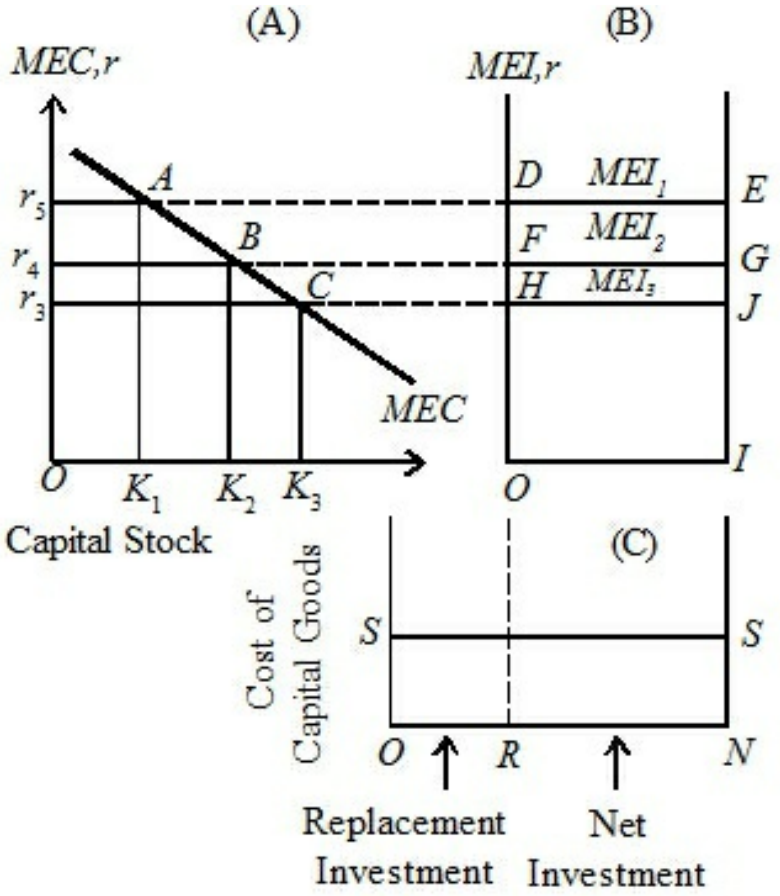


Fig. 7

Panel (A) of the figure shows the MEC curve relating the optimum capital stock to different interest rates. Panel (C) of the figure shows a perfectly elastic supply curve *SS* of the capital goods industry, given the gross investment of *ON*. Out of this *ON* gross investment, the replacement investment is *OR* which is assumed to be constant, as shown by the dotted vertical line from point *R*. Thus the net investment rate per time period available for capital goods industry is *RN*. Panel (B) shows *MEI* curves relating the rate of investment to each market rate of interest. The MEI curves are horizontal (or perfectly elastic) because the supply curve of capital is perfectly elastic.

---

2. A.P. Lerner, *The Economics of Control*, 1946.

3. G. Ackley, *op. cit*., Ch. 17.

Given these, take point *A* on the *MEC* curve in Panel (A) of the figure where the interest rate equals the *MEC*. It is the equilibrium point where the optimum capital stock OK_1 is determined. If the rate of interest falls to

Or_3 and equals the *MEC* at point *C*, OK_3 becomes the optimum or desired capital stock. Now OK_1 is the actual capital stock. So there is a gap between the actual and the desired capital stock equal to K_1K_3. This can be filled by increasing net investment in the capital goods industry. The rate of available net investment is *OI* in Panel (B) or *RN* in Panel (C) which is the capacity level of the capital goods industry.

Investment in capital goods industry can be increased on the basis of the *MEI* of the industry. When the capital stock is OK_1, the *MEI* curve is *DEI* in Panel (B). When investment is made at interest rates below Or_5 the capital stock will increase per time period. Thus in period 1 when net investment is made at *OI* rate, the *MEI* curve is *FGI*, and the capital stock increases to OK_2 from OK_1. In period 2, with the same rate of investment *OI*, when the *MEI* curve is *HJI*, and the capital stock increases to OK_3.* Thus when the rate of interest is Or_3 and equals the *MEC* at point C in Panel (A), the economy attains the optimum level of capital stock where the actual and the desired capital stock are equal. Just as a fall in the interest rate with no shift in the *MEC* curve raises the optimum stock of capital from its initial level, so an upward shift of the *MEC* curve with no change in the rate of interest will have the same result.

If the capital goods industry is operating under the law of increasing costs, its supply curve will be upward sloping which produces the downward sloping *MEI* curve. This is because the cost of capital goods increases as the rate of net investment increases. What will be the rate of net investment in this situation to bring the equilibrium of the actual capital stock with the desired capital stock? This is explained with the help of Figure 8 (A), (B) and (C) Panels which show the *MEC* curve, the *MEI* curve and the rising supply curve of capital respectively.

In Panel (A), the equilibrium point for the capital goods industry is *A* where the interest rate Or_6 and the *MEC* are equal, and OK_1 optimum stock of capital is determined. Here net investment is zero because the actual capital stock equals the desired capital stock of the economy. This is shown by the MEI_1 curve at point *Z* in Panel (B) of the figure. If the rate of interest falls to Or_3, the desired capital stock will be OK_4 and OK_1

becomes the actual capital stock. Thus K_1K_4 is the gap between the actual capital stock and the desired capital stock. To bring equilibrium between the two capital stocks, net investment will have to be made.

---

* The portions *EI*, *GI* and *JI* of *DEI*, *FGI* and *HJI* curves are shown vertical downward.

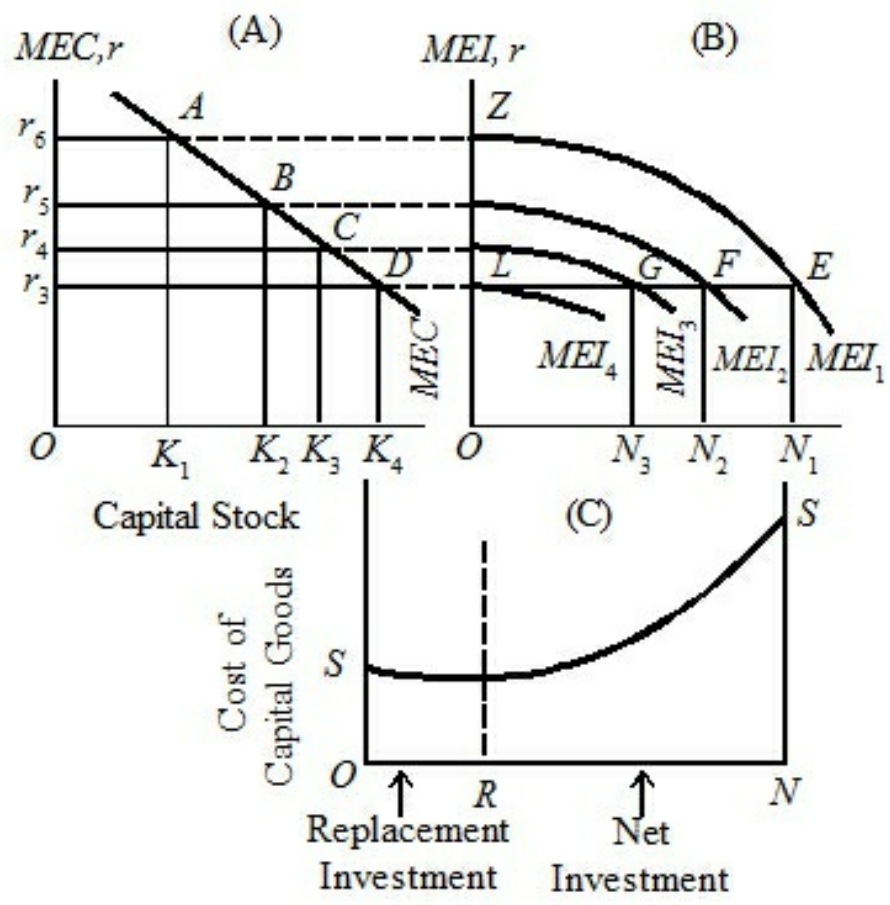


Fig. 8

The rate of net investment will be determined by the equality of the *MEI* with the reduced interest rate Or_3 in each period and so increase the capital stock. Starting from point *Z* on the MEI_1 curve when the rate of interest is Or_6, the rate of investment is zero. As the rate of net investment increases, the prices of capital goods rise and the rate of return on investment in these goods continues to fall till it equals the interest rate Or_3 of the desired capital stock. The MEI_1 curve equals the Or_3 interest rate at point *E* in Panel (B). Thus in period 1, ON_1 rate of net investment increases the capital stock from OK_1 to OK_2, in Panel (A).

This increase in the capital stock by K_1K_2 reduces the *MEC* to point *B* on

the curve *MEC*. Since the rate of interest is Or_3, the capital stock can be increased further. It depends upon the rate of net investment which is determined by the MEI_2 curve at point *F*. So in period 2, the rate of net investment is ON_2 which raises the capital stock to OK_3 from OK_2. This increase in the capital stock by K_2K_3 further reduces the *MEC* to point *C*. The rate of investment in period 3 as determined by the MEI_3 curve corresponding to the level of *MEC* at point *C* is ON_3 which raises the capital stock to the level of MEC at point C is ON_3 which raises the capital stock to reache the optimum level OK_4, the *MEC* equals the rate of interest Or_3 and the corresponding *MEI* curve is MEI_4 which shows zero net investment at point *L*. "The rate of net investment spending per time period depends on how steep the downward slope of the *MEI* curve is (or its elasticity), and this in turn depends on how steep the upward slope (or the elasticity) of the supply curve of capital goods is. If the supply curve slopes sharply upward, the rate of investment spending will fall sharply downward with respect to the rate of interest. In any event, the capital stock will grow to the new *optimum* level, but its rate of growth will be slower the steeper the *MEI* curve."[4]

We have seen above that it is the growth in the capital stock from the actual to the desired level that influences the rate of net investment. This is shown as a downward movement along the *MEC* curve. On the other hand, it is the flow of net investment that adjusts the actual capital stock toward its optimum level along each *MEI* curve with every fall in the rate of interest.

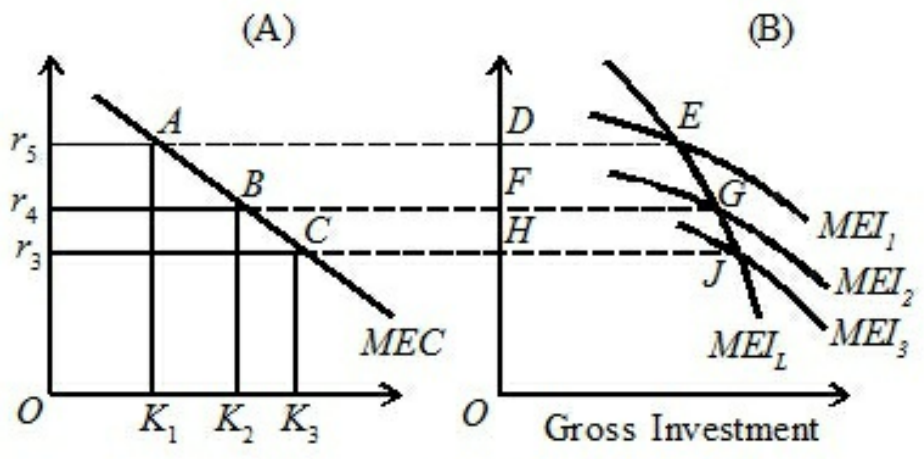


Fig. 9

Now what will be the investment demand in the short-run and the long-run ? Assuming the replacement investment to be given, the increase in the stock of capital and in the rate of net investment relating to fall in the interest rate in each period, as depicted by the *MEC* curve and the MEI_1 curve in Figure 9 (A) and (B), relate to the long-run. If the replacement investment increases with the increase in the capital stock, then the *MEI* curves relating to each level of the rate of interest are the short-run curves, and the long-run *MEI* curve will be as shown in Figure 9. The *MEC* curve, in Panel (A) of the figure relates to the long-run. The curves MEI_1, MEI_2 and MEI_3 are the short-run *MEI* curves. As the rate of interest falls in each period, the capital stock *gradually* increases from OK_1 to OK_2 and finally to OK_3. This is because the rate of replacement investment increases from *DE* to *FG* and to *HJ*. The points *E, G* and *J* where the net investment is zero at each level of interest rate are joined to form the long run *MEI* curve MEI_L.

---

4. Edward Shapiro, *ibid.* Italics mine.

Factors Other Than The Interest Rate Affecting Inducement To Invest

There are a number of factors other than the rate of interest which affect

the inducement to invest. They are the following:[5]

(1) Element of Uncertainty. According to Keynes, the *MEC* is more volatile than the rate of interest. This is because the prospective yield of capital assets depends upon the business expectations. These business expectations are very uncertain. "They may change quickly and drastically in response to the general mood of the business community, rumours, news of technical developments, political events, even directors' ulcers may cause a sudden rise or fall of the expected rate of yield."[6] As a result, it is difficult to calculate the expected annual returns on the life of a capital asset. Further, because of uncertainty, investment projects usually have a short pay-off period. Capital assets become obsolete earlier than their expected life due to rapid technological developments. The rate of depreciation also does not remain constant and varies much. So firms have a tendency to invest only if they are in a position to recover the capital outlay in a short period. These factors tend to bring instability in the investment function.

Harvey and Johnson[7] have pointed out that there is an area of uncertainly around the *MEI* curve. This is due to market imperfections and because the yields from capital assets are based on business expectations. This is expained in Figure 10 where the uncertainty area lies between the two *MEI* curves, MEI_1 and MEI_2. Given the curve MEI_1, a fall in the rate of interest from Or_1 to Or_2 leads to a rise in investment from OI_1 to OI_2. If the MEI_1 curve shifts to MEI_2 investment actually falls to OI_3. However, a shift in the *MEI* curve from MEI_1 to *MEI'* when the rate of interest falls from Or_1 to Or_2 may keep investment at the same level OI_1. Therefore, businessmen are most optimistic around the MEI_1 curve and they are most pessimist around

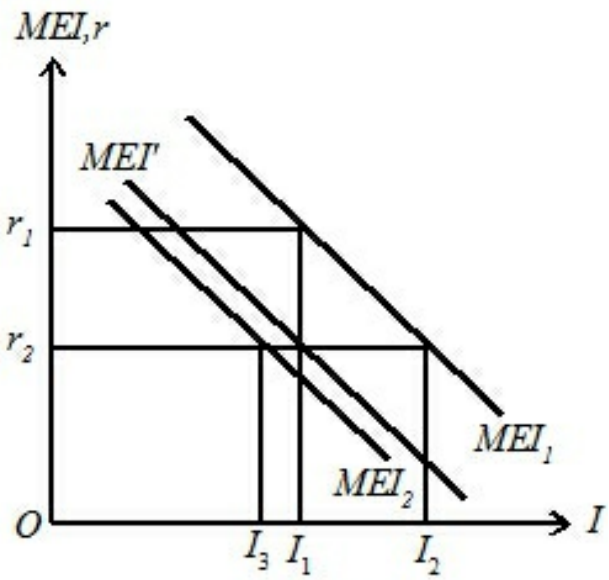


Fig. 10

the MEI_2 curve. To conclude, the *MEI* curve is often so volatile that the rate of interest is a minor influence upon investment decisions.

(2) Existing Stock of Capital Goods. If the existing stock of capital goods is large, it would discourage potential investors from entering into the making of goods. Again, the induced investment will not take place if there is excess or idle capacity in the existing stock of capital assets. In case the existing stock of machines is working to its full capacity, an increase in the demand for goods manufactured by them will raise the demand for capital goods of this type and raise the inducement to invest. But it is the capital stock which influences the *MEC*. The *MEC* and the capital stock are inversely related.

(3) Level of Income. If the level of income rises in the economy through rise in money wage rates the demand for goods will rise which will, in turn, raise the inducement to invest. Contrariwise, the inducement to investment will fall with the lowering of income levels.

---

5. These factors also affect *MEC*. These can be divided into endogenous and exogenous factors. Factors 1 to 5 are endogenous and the remaining factors are exogenous.

6. F.S. Brooman, *Macroeconomics*, 4/e, 1970

7. J.Harvey and M. Johnson, *Introduction to Macroeconomics*, 1971.

(4) Consumer Demand. The present demand and future demand for the products greatly influence the level of investment in the economy. If the current demand for consumer goods is increasing rapidly, more investment will be made. Even if we take the future demand for the products, it will be considerably influenced by their current demand and both will influence the level of investment. Investment will be low if the demand is low, and vice versa.

(5) Liquid Assets. The amount of liquid assets with the investors also influences the inducement to invest. If they possess large liquid assets, the inducement to invest is high. This is especially the case with those firms which keep large reserve funds and undistributed profits. On the contrary, the inducement to invest is low for investors having little liquid assets.

(6) Inventions and Innovations. Inventions and innovations tend to raise the inducement to invest. If inventions and technological improvements lead to more efficient methods of production which reduce costs, the *MEC* of new capital assets will rise. Higher *MEC* will induce firms to make larger investments in the new capital assets and in related ones. The absence of new technologies will mean low inducement to invest. An innovation also includes the opening of new areas. This requires the development of means of transport, the construction of houses, etc. that lead to new investment opportunities. Thus inducement to invest rises.

(7) New Products. The nature of new products in terms of sales and costs may also influence their *MEC* and hence investment. If the sale prospects of a new product are high and the expected revenues more than the costs, the *MEC* will be high which will encourage investment in this and related industries. For example, the invention of television must have encouraged the electronics industry to invest in these capital assets and used them to produce television sets, if they had expected profits to be higher than costs. Thus lower maintenance and operating costs in the case of new products are important in increasing the inducement to invest.

(8) Growth of Population. A rapidly growing population means a growing market for all types of goods in the economy. To meet the demand of an increasing population in all brackets, investment will increase in all types of consumer goods industries. On the other hand, a declining population results in a shrinking market for goods thereby lowering the inducement to invest.

(9) State Policy. The economic policies of the government have an important influence on the inducement to invest in the country. If the state levies heavy progressive taxes on corporations, the inducement to invest is low, and vice versa. Heavy indirect taxation tends to raise the prices of commodities and adversely affects their demand thereby lowering the inducement to invest, and vice versa. If the state follows the policy of nationalisation of industries, the private enterprise would be discouraged to invest. On the other hand, if the state encourages private enterprise by providing credit, power and other facilities, inducement to invest will be high.

(10) Political Climate. Political conditions also affect the inducement to invest. If there is political instability in the country, the inducement to invest may be affected adversely. In the struggle for power, the rival parties may create unrest through hostile trade union activities thereby creating uncertainty in business. On the other hand, a stable government creates confidence in the business community whereby the inducement to invest is raised. Similarly, the danger of a revolution, or war with some other country has an adverse effect on the inducement to invest, whereas peace and prosperity tend to raise it.

EXERCISES

1. Distinguish between (1) Capital and investment, (2) Induced and autonomous investment.
2. What do you mean by investment ? Explain the determinants of the level of investment.
3. What do you mean by marginal efficiency of capital ? How it determines the level of investment ?
4. What do you mean by induced investment ? Discuss the determinants of inducement to invest.
5. Distinguish between marginal efficiency of capital and marginal efficiency of investment.

CHAPTER 12

THE CONCEPT OF MULTIPLIER

INTRODUCTION

The concept of multiplier was first developed by R.F. Kahn in his article "The Relation of Home Investment to Unemployment" in the *Economic Journal* of June 1931. Kahn's multiplier was the Employment Multiplier. Keynes took the idea from Kahn and formulated the Investment Multiplier.

THE INVESTMENT MULTIPLIER

Keynes considers his theory of multiplier as an integral part of his theory of employment. The multiplier, according to Keynes, "establishes a precise relationship, given the propensity to consume, between aggregate employment and income and the rate of investment. It tells us that, when there is an increment of investment, income will increase by an amount which is K times the increment of investment" *i.e.*, $\Delta Y = K\Delta I$. In the words of Hansen, Keynes' investment multiplier is the coefficient relating to an increment of investment to an increment of income, *i.e.*, $K = \Delta Y/\Delta I$, where Y is income, I is investment, Δ is change (increment or decrement) and K is the multiplier.

In the multiplier theory, the important element is the multiplier coefficient, K which refers to the power by which any initial investment expenditure is multiplied to obtain a final increase in income. The value of the multiplier is determined by the marginal propensity to consume. The higher the marginal propensity to consume, the higher is the value of the multiplier,

and vice versa. The relationship between the multiplier and marginal propensity to consume is as follows:

$$\Delta Y = \Delta C + \Delta I$$

or $$\Delta Y = c\Delta Y + \Delta I \qquad [\because \Delta C = c\Delta Y]$$

$$\Delta Y - c\Delta Y = \Delta I$$

$$\Delta Y (1 - c) = \Delta I$$

$$\Delta Y = \frac{\Delta I}{1 - c}$$

$$\frac{\Delta Y}{\Delta I} = \frac{1}{1 - c}$$

$$K = \frac{1}{1 - c} \qquad \left[\because K = \frac{\Delta Y}{\Delta I}\right]$$

Since *c* is the marginal propensity to consume, the multiplier *K* is, by definition, equal to 1–1/*c*. The multiplier can also be derived from the marginal propensity to save(*MPS*) and it is the reciprocal of *MPS*, *K* = 1/*MPS*.

TABLE I: DERIVATION OF THE MULTIPLIER

$\Delta C/\Delta Y(MPC)$	$\Delta S/\Delta Y(MPS)$ $[1-(MPC)]$	K (*Multiplier Coefficient*)
0	1	1
$\frac{1}{2}$	$\frac{1}{2}$	2
$\frac{2}{3}$	$\frac{1}{3}$	3
$\frac{3}{4}$	$\frac{1}{4}$	4
$\frac{4}{5}$	$\frac{1}{5}$	5
$\frac{8}{9}$	$\frac{1}{9}$	9
$\frac{9}{10}$	$\frac{1}{10}$	10
1	0	∞ (Infinity)

The table shows that the size of the multiplier varies directly with the *MPC* and inversely with the *MPS*. Since the *MPC* is always greater than zero and less than one (*i.e.*, $O < MPC <$ I), the multiplier is always between one and infinity (*i.e.*, $1<K<\infty$). If the multiplier is one, it means that the whole increment of income is saved and nothing is spent because the *MPC* is zero. On the other hand, an infinite multiplier implies that *MPC* is equal to one and the entire increment of income is spent on consumption. It will soon lead to full employment in the economy and then create a limitless inflationary spiral. But these are rare phenomena. Therefore, the multiplier coefficient varies between one and infinity.

Working of the Multiplier

The multiplier works both forward and backward. First, we study its forward working. The multiplier theory explains the cumulative effect of a change in investment on income via its efect on consumption expenditure.

Forward Operation

We first take the "sequence analysis" which shows a "motion picture" of the process of income propagation. An increase in investment leads to increased production which creates income and generates consumption expenditure. This process continues in dwindling series till no further increase in income and expenditure is possible. This is a lagless instantaneous process in a static framework, as explained by Keynes.

Suppose that in an economy *MPC* is 1/2 and investment is raised by Rs 100 crores. This will immediately lead to a rise in production and income by Rs 100 crores. One-half of this new income will be immediately spent on consumption goods which will lead to increase in production and income by the same amount, and so on. The process is set out in Table II. It reveals that an increment of Rs 100 crores of investment in the primary round leads to the same increase in income. Of this, Rs 50 crores are saved and Rs 50 crores are spent on consumption which go to increase income by the same amount in the second round. This dwindling process of income generation continues in the secondary rounds till the total income generated from Rs 100 crores of investment rises to Rs 200 crores. This is also clear from the multiplier formula, $\Delta Y = K\Delta I$ or $200=2 \times 100$, where $K = 2$ ($\because MPC = 1/2$) and ΔI = Rs 100 crores.

This process of income propagation as a result of increase in investment is shown diagrammatically in Figure 1.

The *C* curve has a slope of 0.5 to show the *MPC* equal to one-half. $C + I$ is the investment curve which intersects the 45° line at E_1 so that the old equilibrium level of income is OY_1. Now there is and increase in investment of ΔI as shown by the distance between $C + I$ and $C + I + \Delta I$ curves. This curve intersects the 45° line at E_2 to give OY_2 as the new income. Thus the rise in income Y_1Y_2 as shown by ΔY is twice the distance between $C + I$ and $C + I + \Delta I$, since the *MPC* is one-half.

TABLE II: SEQUENCE MULTIPLIER

(Rs Crores)

(Round)	*(Increment in Investment)* ΔI	*(Increment in Income)* ΔY	*(Increment in Consumption)* $\Delta C = c\Delta Y = 0.5$	*(Increment in Saving)* $\Delta S = (\Delta Y - \Delta S)$
0				
1	100	100	50	50
2	—	50	25	25
3	—	25	12.50	12.50
4	—	12.50	6.25	6.25
5	—	6.25	3.12	3.12
0	—	0	0	0
Finally	100	200	100	100

The same results can be obtained if *MPS* is taken so that when income increases, savings also increase to equal the new investment at a new equilibrium level of income. This is shown in Figure 2. *S* is the saving function with a slope of 0.5 to show *MPS* of one-half. *I* is the old investment curve which cuts *S* at E_1 so that OY_1 is the old equilibrium level of income. The increase in investment ΔI is superimposed on the ΔI curve in the shape of a new investment curve $I + \Delta I$ which is intersected by the *S* curve at E_2 to give OY_2 as the new equilibrium level of income. The rise in income $Y_1 – Y_2$ (shown as ΔY)is exactly double the increase in investment ΔI, as the *MPS* is one-half.

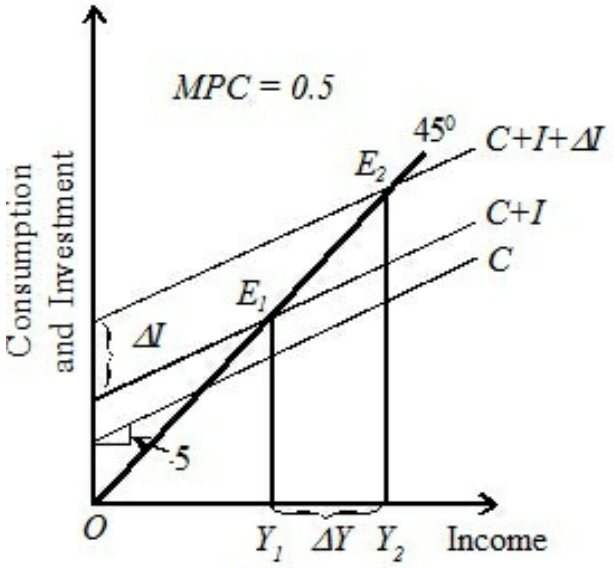


Fig. 1.

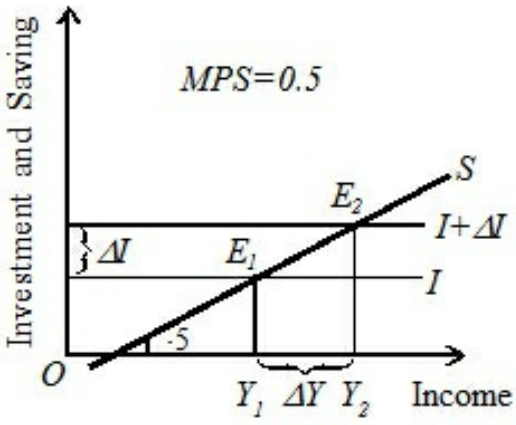


Fig. 2.

Backward Operation

The above analysis pertains to the forward operation of the multiplier. If, however, investment decreases, instead of increasing, the multiplier operates backward. A reduction in investment will lead to contraction of income and consumption which, in turn, will lead to cumulative decline in income and consumption till the contraction in aggregate income is the multiple of the initial decrease in investment. Suppose investment decreases by Rs 100 crores. With an *MPC* = 0.5 and *K*=2, consumption expenditure would keep on declining till aggregate income is decreased by Rs 200 crores. In terms of multiplier formula, $-\Delta Y = K(-\Delta I)$, we get —200 = 2 (–100).

The magnitude of contraction due to the backward operation of the multiplier depends on the value of *MPC*. The higher the *MPC*, the greater is the value of multiplier and the greater the cumulative decline in income, and vice versa. On the contrary, the higher the *MPS*, the lower is the value of the multiplier and the smaller the cumulative decline in income, and vice versa. Thus, a community with a high propensity to consume(or low propensity to save) will be hurt more by the reverse operation of the multiplier than one with a low propensity to consume(or high propensity to save).

Diagrammatically, the reverse operation also can be explained in terms of Figures 1. and 2. Taking Figure 1., when investment decreases, the investment function $C + I + \Delta I$ shifts downward to $C + I$. As a result, the equilibrium level also shifts from E_2 to E_1 and income declines from OY_2 to OY_1. The *MPC* being 0.5, the fall in income Y_2Y_1 is exactly double the decline in investment as shown by the distance between $C + I + \Delta I$ and $C + I$. Similarly, in Figure 2 when investment falls, the investment function $I + \Delta I$ shifts downward as I curve and income decreases from OY_2 to OY_1. The *MPS* being 0.5, the decrease in income Y_2Y_1 is double the decline in investment as measured by the distance between $I + \Delta I$ and I curves.

Assumptions of Multiplier

Keynes's theory of the multiplier works under certain assumptions which limit the operation of the multiplier. They are as follows:

(1) There is change in autonomous investment and that induced investment is absent.

(2) The marginal propensity to consume is constant.

(3) Consumption is a function of current income.

(4) There are no time lags in the multiplier process. An increase(decrease) in investment instantaneouly leads to a multiple increase(decrease) in income.

(5) The new level of investment is maintained steadily for the completion of the multiplier process.

(6) There is net increase in investment.

(7) Consumer goods are available in response to effective demand for them.

(8) There is surplus capacity in consumer goods industries to meet the increased demand for consumer goods in response to a rise in income following increased investment.

(9) Other resources of production are also easily available within the economy.

(10) There is an industrialised economy in which the multiplier process operates.

(11) There is a closed economy unaffected by foreign influences.

(12) There are no changes in prices.

(13) The accelerator effect of consumption on investment is ignored.

(14) There is less than full employment level in the economy.

Leakages of Multiplier

Leakages are the potential diversions from the income stream which tend to weaken the multiplier effect of new investment. Given the marginal propensity to consume, the increase in income in each round declines due to leakages in the income stream and ultimately the process of income propagation "peters out." (see Table II). The following are the important leakages:

1. Saving. Saving is the most important leakage of the multiplier process. Since the marginal propensity to consume is less than one, the whole increment in income is not spent on consumption. A part of it is saved which peters out of the income stream and the increase in income in the next round declines. Thus the higher the marginal propensity to save, the smaller the size of the multiplier and the greater the amount of leakage out of the income stream, and vice versa. For instance, if *MPS* = 1/6, the multiplier is 6, according to the formula $K=1/MPS$; and the *MPS* of 1/3 gives a multiplier of 3.

2. Strong Liquidity Preference. If people prefer to hoard the increased income in the form of idle cash balances to satisfy a strong liquidity preference for the transaction, precautionary and speculative motives, that will act as a leakage out of the income stream. As income increases people will hoard money in inactive bank deposits and the multiplier process is

checked.

3. Purchase of Old Stocks and Securities. If a part of the increased income is used in buying old stocks and securities instead of consumer goods, the consumption expenditure will fall and its cumulative effect on income will be less than before. In other words, the size of the multiplier will fall with a fall in consumption expenditure when people buy old stocks and shares.

4. Debt Cancellation. If a part of increased income is used to repay debts to banks, instead of spending it for further consumption, that part of the income peters out of the income stream. In case, this part of the increased income is repaid to other creditors who save or hoard it, the multiplier process will be arrested.

5. Price Inflation. When increased investment leads to price inflation, the multiplier effect of increased income may be dissipated on higher prices. A rise in the prices of consumption goods implies increased expenditure on them. As a result, increased income is absorbed by higher prices and the real consumption and income fall. Thus price inflation is an important leakage which tends to dissipate increase in income and consumption on higher prices rather than in increasing output and employment.

6. Net Imports. If increased income is spent on the purchase of imported goods it acts as a leakage out of the domestic income stream. Such an expenditure fails to effect the consumption of domestic goods. This argument can be extended to *net* imports when there is an excess of imports over exports thereby causing a net outflow of funds to other countries.

7. Undistributed Profits. If profits accruing to joint stock companies are not distributed to the shareholders in the form of dividend but are kept in the reserve fund, it is a leakage from the income stream. Undistributed profits with the companies tend to reduce the income and hence further expenditure on consumption goods thereby weakening the multiplier process.

8. Taxation. Taxation policy is also an important factor in weakening the

multiplier process. Progressive taxes have the effect of lowering the disposable income of the taxpayers and reducing their consumption expenditure. Similarly commodity taxation tends to raise the prices of goods, and a part of increased income may be dissipated on higher prices. Thus increased taxation reduces the income stream and lowers the size of the multiplier.

9. Excess Stocks of Consumption Goods. If the increased demand for consumption goods is met from the existing excess stocks of consumption goods there will be no further increase in output, employment and income and the multiplier process will come to a halt till the old stocks are exhausted.

10. Public Investment Programmes. If the increase in income as a result of increased investment is affected by public expenditures, it may fail to induce private enterprise to spend that income for further investment due to the following reasons.

(a) Public investment programmes may raise the demand for labour and materials leading to a rise in the costs of construction so as to make the undertaking of some private projects unprofitable.

(b) Government borrowing may, if not accompanied by a sufficiently liberal credit policy on the part of the monetary authority, increase the rate of interest and thus discourage private investment.

(c) Government operations may also injure private investors' confidence by arousing animosity or fears of nationalisation.

Criticism of Multiplier

The multiplier theory has been severely criticised by the post-Keynesian economists on the following grounds:

1. Merely Tautological Concept. Prof. Haberler has criticised Keynes' multiplier as tautological. It is a truism which defines the multiplier as

necessarily true as $K = \frac{1}{1 - \frac{\Delta C}{\Delta Y}}$. As pointed by Professor Hansen, "Such a coefficient is a mere *arithmetic* multiplier(*i.e.*, a truism) and not a true *behaviour* multiplier based on a *behaviour* pattern which establishes a verifiable relation between consumption and income. A mere arithmetic multiplier, $\frac{1}{1 - \frac{\Delta C}{\Delta Y}}$, is tautological."

2. Timeless Analysis. Keynes's logical theory of the multiplier is an instantaneous process without time lag. It is a timeless static equilibrium analysis in which the total effect of a change in investment on income is instantaneous so that consumption goods are produced simultaneously and consumption expenditure is also incurred instantaneously. But this is not borne out by facts because a time lag is always involved between the receipt of income and its expenditure on consumption goods and also in producing consumption goods. Thus "the timeless multiplier analysis disregards the transition and deals only with the new equilibrium income level" and is therefore unrealistic.

3. Worthless Theoretical Toy. According to Hazlitt, the Keynesian multiplier "is a strange concept about which some Keynesians make more fuss than about anything else in the Keynesian system." It is a myth for there can never by any precise, predeterminable or mechanical relationship between investment and income. Thus he regards it as "a worthless theoretical toy."

4. Acceleration Effect Ignored. One of the weaknesses of the multiplier theory is that it studies the effects of investment on income through changes in consumption expenditure. But it ignores the effect of consumption on investment which is known as the acceleration principle. Hicks, Samuelson and others have shown that it is the interaction of the multiplier and the accelerator which helps in controlling business

fluctuations.

5. MPC does not Remain Constant. Gordon points out that the greatest weakness of the multiplier concept is its exclusive emphasis on consumption. He favours the use of the temn 'marginal propensity to spend' in place of marginal propensity to consume to make this concept more realistic. He also objects to the constancy of the marginal propensity to spend (or consume) because in a dynamic economy, it is not likely to remain constant. If it is assumed to be constant, it is not possible "to predict with much accuracy the multiplying effects over the cycle of a given increase in private investment or public spending."

6. Relation between Consumption and Income. Keynes's multiplier theory establishes a linear relation between consumption and income with the hypothesis that the *MPC* is less than one and greater than zero. Empirical studies of the behaviour of consumption in relation to income show that the relationship between the two is complicated and non-linear. As pointed out by Gardner Ackley, "The relationship does not run simply from *current* income to *current* consumption, but rather involves some complex average of past and expected income and consumption. There are other factors than income to consider."

Other economists have not been lagging behind in their criticism of the multiplier concept. Prof. Hart considers it "a useless fifth wheel." To Stigler, it is the fuzziest part of Keynes's theory. Prof. Hutt calls it a "rubbish apparatus" which should be expunged from text books.

But despite its scathing criticism, the multiplier principle has considerable practical applicability to economic problems as given below.

Importance of Multiplier

The concept of mutiplier is one of the important contributions of Keynes's to the income and employment theory. As aptly observed by Richard Goodwin. "Lord Keynes did not discover the multiplier; that honour goes to Mr R.F. Kahn. But he gave it the role it plays today by transforming it from an instrument for the analysis of road building into one for the analysis of income building....It set a fresh wind blowing through the

structure of economic thought." Its importance lies in the following:

1. Investment. The multiplier theory highlights the importance of investment in income and employment theory. Since the consumption function is stable during the short-run fluctuations in income and employment are due to fluctuations in the rate of investment. A fall in investment leads to a cumulative decline in income and employment by the multiplier process and *vice versa*. Thus it underlines the importance of investment and explains the process of income propagation.

2. Trade Cycle. As a corollary to the above, when there are fluctuations in the level of income and employment due to variations in the rate of investment, the multiplier process throws a spotlight on the different phases of the trade cycle. When there is a fall in investment, income and employment decline in a cumulative manner leading to recession and ultimately to depression. On the contrary, an increase in investment leads to revival and, if this process continues, to a boom. Thus the multiplier is regarded as an indispensable tool in trade cycles.

3. Saving-Investment Equality. It also helps in bringing the equality between saving and investment. If there is a divergence between saving and investment, and increase in investment leads to a rise in income via the multiplier process by more than the increase in initial investment. As a result of the increase in income, saving also increases and equals investment.

4. Formulation of Economic Policies. The multiplier is an important tool in the hands of modern states in formulating economic policies. Thus this principle pre-supposes state intervention in economic affairs.

(a) To achieve full employment. The state decides upon the amount of investment to be injected into the economy to remove unemployment and achieve full employment. An initial increase in investment leads to the rise in income and employment by the multiplier time the increase in investment. If a single dose of investment is insufficient to bring full employment, the state can inject regular doses of investment for this purpose till the full employment level is reached.

(b) To control trade cycles. The state can control booms and depressions in a trade cycle on the basis of the multiplier effect on income and employment. When the economy is experiencing inflationary pressures, the state can control them by a reduction in investment which leads to a cumulative decline in income and employment via the multiplier process. On the other hand, in a deflationary situation, an increase in investment can help increase the level of income and employment through the multiplier process.

(c) Deficit financing. The multiplier principle highlights the importance of deficit budgeting. In a state of depression, cheap money policy of lowering the rate of interest is not helpful because the marginal efficiency of capital is so low that a low rate of interest fails to encourage private investment. In such a situation, increased public expenditure through public investment programmes by creating a budget deficit helps in increasing income and employment by multiplier time the increase in investment.

(d) Public investment. The above discussion reveals the importance of the multiplier in public investment policy. Public investment refers to the state expenditure on public works and other works meant to increase public welfare. It is autonomous and is free from profit motive. It, therefore, applies with greater force in overcoming inflationary, and deflationary pressures in the economy, and in achieving and maintaining full employment. Private investment being induced by profit motive can help only when the public investment has created a favourable situation for the former. Moreover, economic activity cannot be left to the vagaries and uncertainties of private enterprise. Hence, the importance of multiplier in public investment lies in creating or controlling income and employment. The state can have the greatest multiplier effect on income and employment by increasing public investment during a depression where the *MPC* is high (or the *MPS* is low). On the contrary, in periods of overfull employment, a decline in investment will have a serious effect on the levels of income and employment where the *MPS* is high (or *MPC* is low). The best policy is to reduce investment where the *MPC* is low(or *MPS* is high), to have gradual decline in income and employment. The important thing, however, is the timing of public investment in such a

manner that the multiplier is able to work with full force and there is little scope for the income stream to peter out. Moreover, public investment should not supplant but supplement private investment so that it could be increased during depression and reduced during inflation. As a result, the forward and backward operation of the multiplier will help in the two situations.

THE DYNAMIC OR PERIOD MULTIPLIER

Keynes's logical theory of the multiplier is an instantaneous process *without time lags*. It is a timeless static equilibrium analysis in which the total effect of a change in investment on income is instantaneous so that consumption goods are produced simultaneously and consumption expenditure is also incurred instantaneously. But this is not borne out by facts because a time lag is always involved between the receipt of income and its expenditure on consumption goods and also in producing consumption goods. Thus "the timeless multiplier analysis disregards the transition and deals only with the new equilibrium income level" and is, therefore, unrealistic.

The dynamic multiplier relates to the time lags in the process of income generation. The series of adjustments in income and consumption may take months or even years for the multiplier process to complete, depending upon the assumption made about the period involved. This is explained in Table III where if each round is of one month and it takes seventeen rounds for an initial investment of Rs 100 crores to generate an income of Rs 200 crores, given the value of *MPC* to be 0.5, then the multiplier process will take 17 months to complete.

TABLE III: DYNAMIC OR PERIOD MULTIPLIER

(Rs. Cror

Period in Months	*ΔI (Increment in investment)*	*ΔC = cΔY = 0.5 Increment in consumption)*	*ΔY (Increment in income)*
0	0	0	0
t+1	100	0	100

t+2	100	50	100+50
t+3	100	25	150+25
:	:	:	:
t+n	100	100	200

The Table shows that if the *MPC* remains constant at 0.5 throughout, an initial increase of Rs 100 crores of investment will first raise income by 100 crores in the first month. Out of this Rs 50 crores will be spent on consumption. This will raise income in the second month to Rs 50 crores, and out of this Rs 25 crores will be spent on consumption. This will go to increase income in the third month by Rs 25 crores, and successive increments in income get smaller and smaller in each period till in the seventeenth month the income increases by Rs 0.001 crore. This can also be explained algebraically as:

$\Delta Y = \Delta I + \Delta Ic + \Delta Ic^2 + \Delta Ic^3 + ... + \Delta Ic^{n-1}$ [*c is MPC*]

$\Delta Y = 100 + 100\ (0.5) + 100\ (0.5)^2 + 100\ (0.5)^3 + ...100\ (0.5)^{n-1}$

$$\Delta Y = \frac{1-.5^2}{1-.5}100 = \frac{1}{1-.5}100 = \text{Rs 200 crores.}$$

This process of dynamic income propagation assumes that there is a consumption lag and no investment lag so that consumption is a function of the income of the preceding period *i.e.*, $C_t = f\ (Y_{t-1})$ and investment is a function of time (t) and of constant autonomous investment , ΔI, *i.e.* $I_t = f\ (\Delta I)$. In Figure 3, $C + I$ is the aggregate demand function and the 45° line is the aggregate supply function. If we begin in period t_O where with an equilibrium level of OY_O income, investment is increased by ΔI, then in period t income rises by the amount of the increased investment (from $t_O tc\ t$). The increased investment is shown by the new aggregate demand function $C + I + \Delta I$. But in period t_O consumption lags behing, and is still equal to the original income E_O. But at Y_O level total demand rises from $Y_O t_O$ to $Y_O t$. There is now an excess of demand over supply equal to $t_O t$. In period t consumption rises due to

the rise in demand to $Y_O t$. Now investment increases income still higher to OY_1 in period t +1 and to increase in consumption from t to E_1. But at this level, total demand is Y_1E_1 which exceeds total supply by AE_1. This will further tend to raise income to OY_2 in period t+2 and to increase in consumption to E_1E_2. This leads to a rise in demand to Y_2E_2, leading to an excess of total demand over total supply by BE_2. This process of income generation will continue till the aggregate demand function $C + I + \Delta I$ equals the aggregate supply function 45° line at E_n in the *nth* period, and the new equilibrium level of income is determined at OY_n. The curved steps E_0 to E_n is the path of income propagation showing the dynamic process of multiplier. The lower portion of the figure shows the time dimension of the multiplier process.

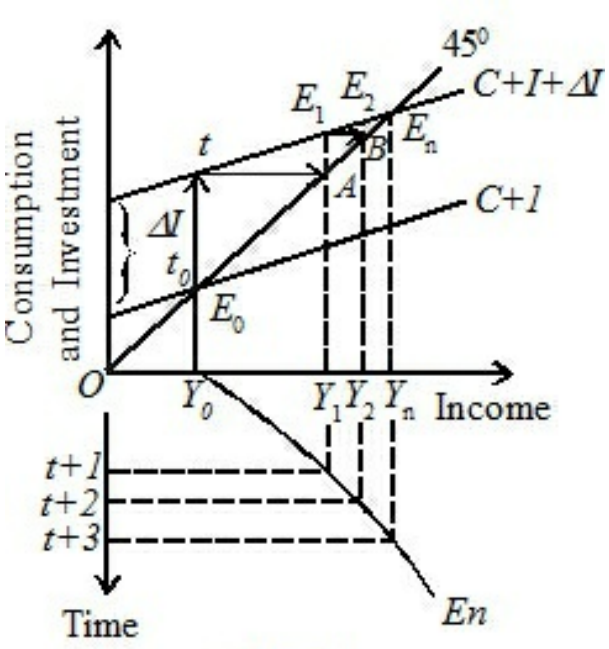


Fig. 3.

THE EMPLOYMENT MULTIPLIER

The concept of Employment Multiplier was introduced by R.F. Kahn in 1931 as a ratio between the total increase in employment and primary employment, *i.e.* $K^1=\Delta N/\Delta N_1$ where K^1 stands for the employment multiplier, ΔN for the increase in total employment and ΔN_1 for the increase in primary employment. Thus the "employment multiplier is a coefficient relting an increment of primary employment on public works to the resulting increment of total employment, primary and secondary combined." To illustrate it, suppose 200000 additional men are employed in public works so that the (secondary) employment is increased by 400000. The toal employment is increased by 600000 (=200000 primary + 400000 secondary). The employment multiplier would be

600000/200000=3.

Algebraically, the Keynesian multiplier $\Delta Y = K\Delta I$ is analogous to Kahn's multiplier $\Delta N = K^1 \Delta N_1$. But Keynes points out that there is no reason in general to suppose that $K = K^1$ because income in terms of wage units may rise more than employment, if in the process, nonwage earners' income should rise proportionately more than wage earners' income. Moreover, with decreasing returns, total product would rise proportionately less than employment. In short, income in terms of wage units would rise most, employment next and output the least. Still, according to Hansen, in the short-run, all three would tend to rise and fall together as envisaged by the Keynesian income and employment theory. He concludes that thus for practical purposes we do no great violence to the facts if we assume that the employment multiplier K^1 equals the investment multiplier K.

If, however, output increases towards the full employment output, per unit of labour will fall due to decreasing returns. In such a situation, K^1 is larger than K when the multiplier is working to increase ouput and employment. But K^1 is smaller than K if the multiplier is working in the opposite direction.

Dillard points out the employment multiplier is useful for showing the relation between primary and secondary employment from public works. But Keynes' conception is superior to Kahn's because in the words of Goodwin, "He gave it the role it plays today by transforming it from an instrument for the analysis of road building into one for the analysis of income building."

SELECTED READINGS

F.S. Brooman, *Macro Economics*, 4/e, 1970, Ch,VI.

J.H. Makwin, *Macroeconomics*, 1975, Ch. 5.

Edward Shapiro, *Macroeconomic Analysis*, 4/e, 1978, Ch.5.

P. Wonnacott, *Macroeconomics*, 1974, Ch. 3.

A.H. Hansen, *A Guide to Keynes*, 1953, Ch.4.

EXERCISES

1. What is the significance of the statement that the value of the multiplier is the reciprocal of the marginal propensity to save ?
2. Discuss the concept of "investment multiplier" and its role in the theory of income and employment.
3. Discuss the concept of multiplier as put forward by Kahn and Keynes.
4. For what analytical purposes are the notions of "leakages" and "lags" employed in the theory of income propagation ?
5. Define the "Multiplier". Show the way the multiplier has been used in the Keynesian theory of employment.
6. "The higher the marginal propensity to consume, the higher will be the value of the multiplier." Discuss this statement.
7. Explain the dynamic multiplier. How is it different from Keynes' multiplier ?

CHAPTER 13

COMPLEX MULTIPLIERS

INTRODUCTION

Keynes' investment multiplier is simple and static in which income depends upon consumption and investment. It is called a two sector model. After Keynes, in order to make the multiplier more practical, economists included a number of variables to construct many multipliers which are called complex multipliers. These are dynamic multiplier, government expenditure multiplier, tax multipliers, balanced budget multiplier and foreign trade multiplier.

Keynes' two sector model depends upon consumption and investment. By including government expenditure and taxes, it becames a three sector model. When exports and imports are included in it, it becomes a four sector model. These sector models are discussed in the chapter on Income Determination in a Closed and Open Economy.

The present chapter explains government expenditure multiplier, tax multipliers and balance budget multiplier. The foreign trade multiplier is discussed in the next chapter.

GOVERNMENT EXPENDITURE MULTIPLIER

The Keynesian investment multiplier is in fact expenditure multiplier which measures the rate of change in income due to a change in autonomous consumption expenditure and autonomous investment

expenditure, $K = \frac{1}{1-c}$.

Similarly, government expenditure multiplier *K*g is a change in income due to a change in autonomous government expenditure. It can be expressed as:

$$\Delta Y = \frac{1}{1-c}\Delta G$$

$$\frac{\Delta Y}{\Delta G} = \frac{1}{1-c}$$

$$Kg = \frac{\Delta Y}{\Delta G} = \frac{1}{1-c}$$

which shows that change in income (ΔY) is equal to the multiplier (1/1-c) multiplied by change in government autonomous expenditure (ΔG). If c = 2/3, then $Kg = \frac{1}{1-2/3} = 3$ which is the value of government expenditure multiplier.

The government expenditure multiplier is shown in Fig. 1 where income is taken on the horizontal axis and government expenditure (*C+I+G*) is taken on the vertical axis. According to Keynes' two sector model, *C+I* is the total expenditure curve which cuts the 45° curve at point *E* and *OY* is the initial equilibrium income level. By adding government expenditurre (G), the *C+I* curve shifts upward and becomes *C+I+G* curve which

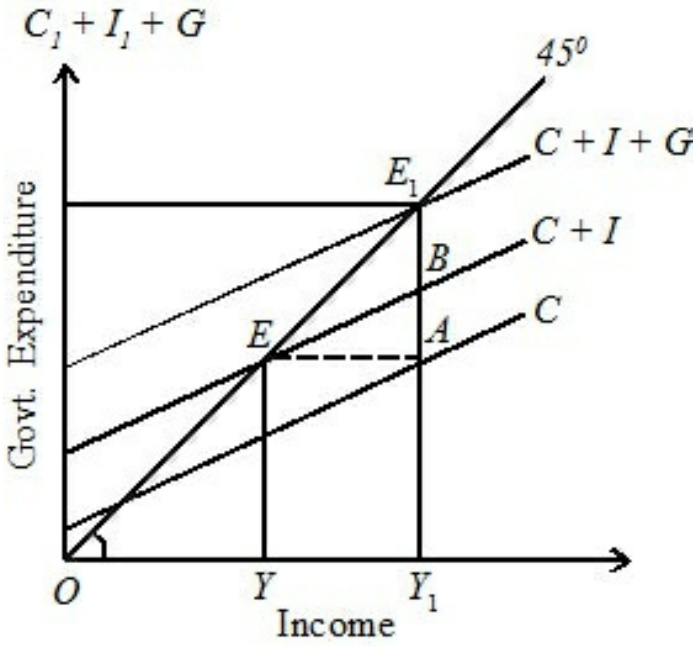


Fig. 1

intersects the 45° line at point E_1. Now OY_1 is the new equilibrium level of income. As a result of the government expenditure multiplier, the increase in income YY_1 (=*EA*) is more than the government expenditure BE_1. This shows that the government expenditure multiplier is more than unity, as 3 in our above example.

TAX MULTIPLIERS

When the government changes the tax rates, the relation between disposable income and national income changes. When the government increases a tax rate (*T*) or levies a new tax, the marginal propensity to consume (*c*) of the people declines because their disposal income is reduced. This brings a fall in national income due to the multiplier effect. On the other hand, reduction in taxes has the multiplier effect of raising the national income. The tax multiplier (K_T) is

$$\Delta Y = \frac{-c}{1-c}\Delta T$$

$$K_T = \frac{\Delta Y}{\Delta T} = \frac{-c}{1-c}$$

$$\text{If } c = 2/3, \text{ then } K_T = \frac{-2/3}{1-2/3} = -2$$

Government usually levies two types of taxes, lumpsum and proportional.

First, we explain *lumpsum tax multiplier* in Fig. 2. Before the levy of a lumpsum tax, *C* is the consumption function and the income level is *OY*. Now AG amount of tax is levied. As a result, the disposable income is reduced and the consumption function shifts downward from *C* to C_1. With the decline in the consumption function, the total expenditure curve (*C*+*I*+*G*) also shifts downward to *C*+*I*+*G*–*T* curve. This intersects the 45is the total expenditure curve which cuts the 45° line at E_1 and the national income is reduced from *OY* to OY_1.

Second, if the government levies a *proportional income tax*, this also brings a fall in the consumption function due to a decline in disposable income of the people. Consequently, the national income declines due to the tax multiplier. This is shown is Fig. 3, where *C* is the consumption function before the tax is levied and *OY* is the income level. When *AT* tax is levied, the *C* curve revolves downward to C_1. With the fall in the consumption function, the total expenditure curve (*C*+*I*+*G*) also revolves downward to *C*+*I*+*G*–*T* and intersects the 45° line at E_1. This brings reduction in national income from *OY* to OY_1 .

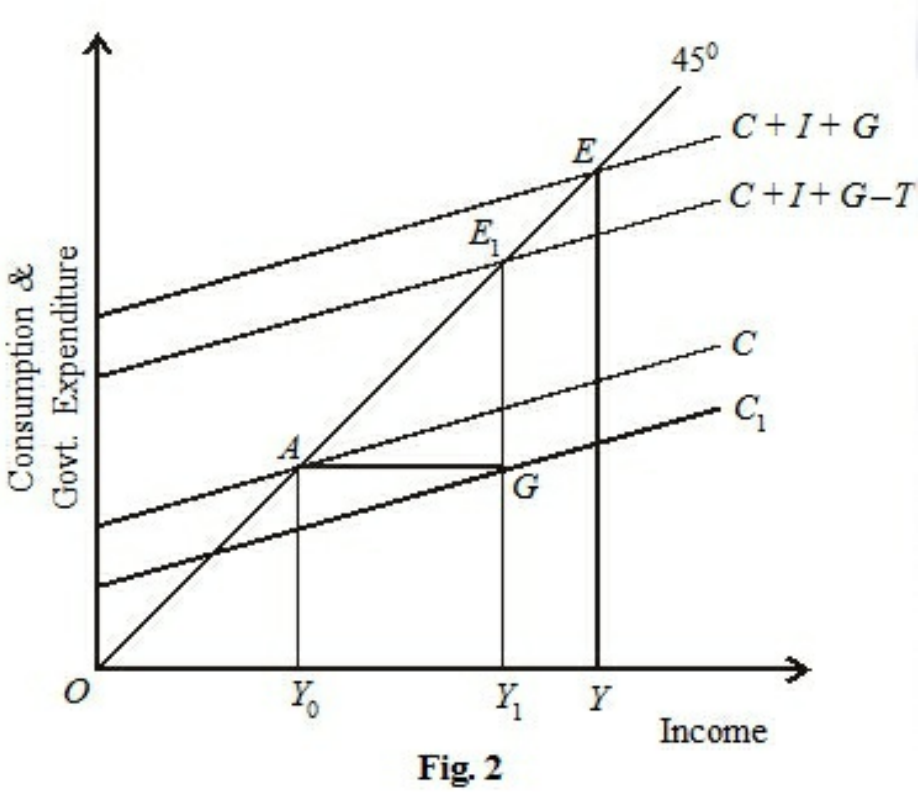


Fig. 2

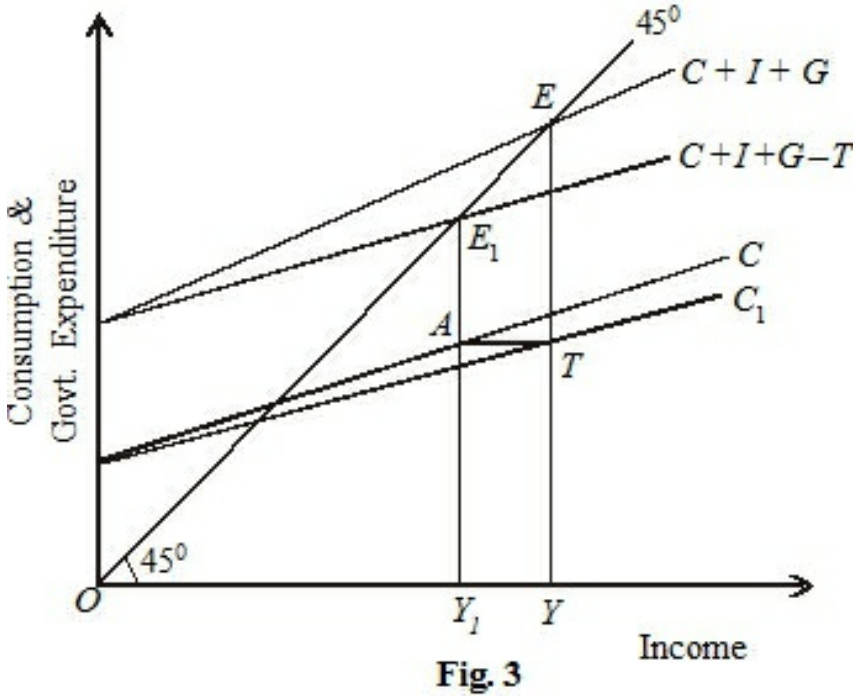


Fig. 3

BALANCED BUDGET MULTIPLIER

The balanced budget multiplier is used to show an expansionist fiscal policy. In this the increase in taxes (ΔT) and in government expenditure (ΔG) are of an equal amount ($\Delta T=\Delta G$). Still there is increase in income. The basis for the expansionary effect of this kind of balanced budget is that a tax merely tends to reduce the level of disposable income. Therefore, when only a portion of an economy's disposable income is used for consumption purposes, the economy's consumption *expenditure* will not fall by the full amount of the tax. On the other hand, government *expenditure* increases by the full amount of the tax. Thus the government expenditure rises more than the fall in consumption expenditure due to the tax and there is net increase in national income.

The balanced budget multiplier is based on the combined operation of the tax multiplier and the government expenditure multiplier. In the balanced budget multiplier, the tax multiplier is smaller than the government expenditure multiplier. The government expenditure multiplier is

$$\Delta Y = \frac{1}{1-c}\Delta G$$

$$\text{or} \quad Kg = \frac{\Delta Y}{\Delta G} = \frac{1}{1-c} \qquad \text{...(1)}$$

which indicates that the change in income (ΔY) will equal the multiplier (1/1–c) *times* the change in autonomous government expenditure.

The tax multiplier is

$$\Delta Y = \frac{-c\Delta T}{1-c}$$

$$\text{or} \quad K_T = \frac{\Delta Y}{\Delta T} = \frac{-c}{1-c} \qquad \text{...(2)}$$

which shows that the change in income (ΔY) will equal multiplier (1/1–c) *times* the product of the marginal propensity to consume (c) and the

change in taxes (ΔT).

A simultaneous change in public expenditure and taxes may be expressed as a combination of equations (1) and (2) which is balanced budget multiplier,

$$K_b = \frac{\Delta Y}{\Delta G} + \frac{\Delta Y}{\Delta T} = \frac{1}{1-c} + \frac{-c}{1-c} = \frac{1-c}{1-c} = 1$$

or $K_b = 1$

Since $\Delta G = \Delta T$, income will change (ΔY) by an amount equal to the change in government expenditure (ΔG) and taxes (ΔT).

To understand it, it is explained numerically. Suppose the value of $c = 2/3$ and the increase in government expenditure ΔG = Rs 10 crores. Since $\Delta G = \Delta T$, therefore the increase in lumpsum taxes ΔT = Rs 10 crores.

We first calculate the government expenditure multiplier

$$Kg = \frac{\Delta Y}{\Delta G} = \frac{1}{1-c} = \frac{1}{1-2/3} = 3$$

$$\text{Tax multiplier, } K_T = \frac{\Delta Y}{\Delta T} = \frac{-c}{1-c} = \frac{-2/3}{1-2/3} = -2$$

To arrive at the increase in income as a result of the combined operation of the government expenditure multiplier and the tax multiplier, we write the balanced budget multiplier equation as

Thus the increase in income (ΔY) exactly equals the increase in government expenditure (ΔG) and the lumpsum tax (ΔT) i.e. Rs. 10 crores. Thus K_b=1.

$$K_b = \Delta Y = \frac{1}{1-c}\Delta G - \frac{-c}{1-c}\Delta T$$

and fit in the above values of c, ΔG and ΔT so that

$$K_b = \Delta Y = 3\Delta G - 2\Delta T$$
$$= 3\times 10 - 2\times 10$$
$$= \text{Rs. } 10 \text{ crores.}$$

This balanced budget multiplier or unit multiplier is explained in Fig. 4. C is the consumption function before the imposition of the tax with income at OY_0 level. Tax of AG amount is imposed. As a result, the consumption function shifts downward to C_1. Now government expenditure of GE amount is injected into the economy which is equal to the tax yield AG. The new government expenditure line is C_1+G which determines OY income at point E. The increase in income $Y_0\ Y$ equals the tax yield AG and the increase in government expenditure GE. This proves that income has risen by 1 (one) *time* the amount of increase in government expenditure which is a balanced budget expansion. This analysis relates to the imposition of a *lumpsum tax*. However, when a *lumpsum tax* is levied, the *MPC* of national income is reduced, and the

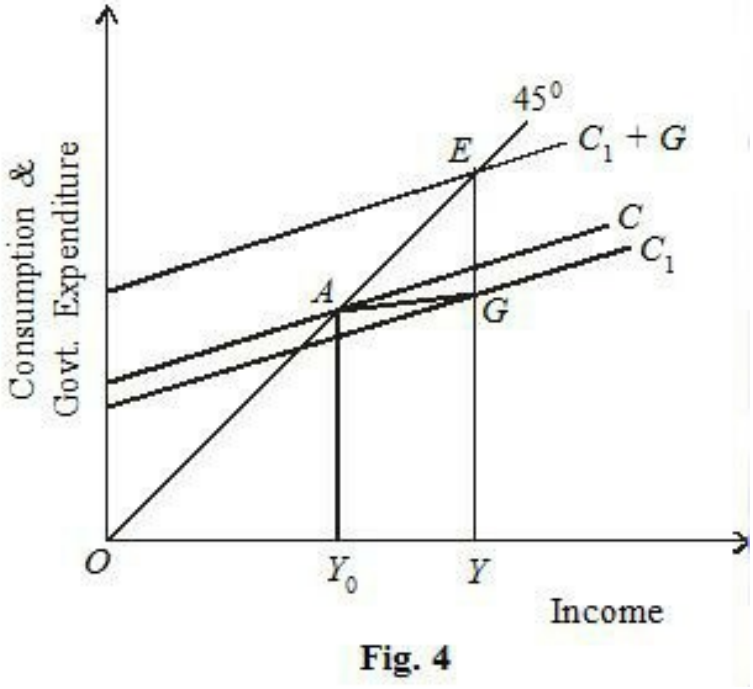


Fig. 4

value of the multiplier is less than under the lumpsum tax. The multiplier formula in this case is $\frac{\Delta Y}{\Delta G} = \frac{1}{1-c(1-t)}$. Here the term c (1–t) is the *MPC* of taxable national income. Thus the fraction of taxable national income spent on consumption will equal $c(1–t)$. In this case, an increase in government expenditure raises the disposable income only by (1–t) *times* the increase in income because a proportion of the tax levied (t) goes to the government exchequer. Consequently, the *MPC* of national income is reduced and the value of the multiplier is low, as per the above equation. This can be explained with the help of an example.

Suppose the tax rate (t) = 25%. Thus (1–t) = 1–1/4 and by assuming the value of c (MPC) = 2/3, the government expenditure multiplier with lumpsum tax is

$$\frac{\Delta Y}{\Delta G} = \frac{1}{1-c(1-t)} = \frac{1}{1-2/3(1-1/4)} = \frac{1}{1/2} = 2$$

which is less than the government expenditure multiplier without a tax, i.e.,

$$\frac{\Delta Y}{\Delta G} = \frac{1}{1-c} = \frac{1}{1-2/3} = \frac{1}{1/3} = 3$$

This analysis shows that when a lumpsum income tax is levied the disposable income level is reduced and a portion of the government's increased income due to tax collection goes to the exchequer. Thus the expansionary effect of the government expenditure becomes ineffective and the balanced budget multiplier operates.

But when a *proportional income tax* is levied, the government expenditure is increased by the full amount of the tax revenue, and nothing goes to the exchequer, the balanced budget theorem holds. This is illustrated in Fig. 5 where C is the consumption function before the imposition of the income

tax. An income tax equal to Y_1Y_2 /OY_2 is levied. As a result, the old consumption function pivots to the lower position of C_1. The tax revenue going to the exchequer is *AG*. Now government expenditure is equal to the tax revenue. This is *GE* = *AG* which is injected into the economy. The new government expenditure line C_1+*G* determines OY_2 national income at point *E*. The increase in income Y_1Y_2 equals the tax revenue *AG* and the increase in government expenditure *GE*. Thus the increase in income exactly equals the increase in the tax revenue and the government expenditure. This proves the balanced budget theorem under *proportional income tax*. The analysis also shows that even after the imposition of income tax, there is no reduction in the *MPC* of individuals. It remains unchanged AY_1=GY_2.

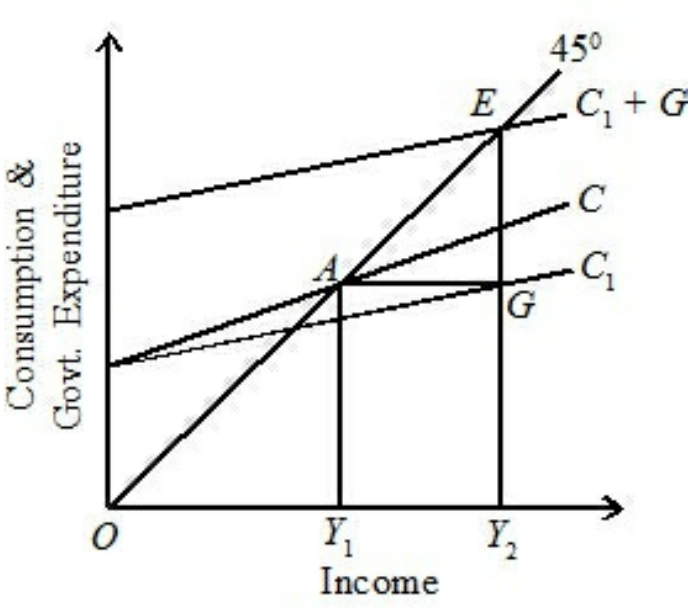


Fig. 5

But this is highy unrealistic because the tax rate increases and lowers the level of disposable income and the government is not able to match its expenditure equal to the tax yield.

Its Limitations

The concept of balanced budget multiplier has the following limitations:

1. It takes into account only government expenditure on goods and services and excludes transfer payments. In fact a transfer payment's multiplier offsets the negative tax multiplier.

2. It assumes a uniform *MPC* for those who pay taxes and those who sell their goods and services to the government.

3. It does not take into consideration the impact of government

expenditure and taxes on investment. So far as taxes are concerned, they affect either investment or consumption depending upon the type of tax-payers, whether the tax is levied on the business community or the fixed income groups.

Its Critical Appraisal

Besides the above limitations, the use of the balanced budget as an expansionary device has been found inefficient and inadequte. This policy requires large government expenditure which may lead to a considerable diversion in the allocation of resources from the private to the public sector, thereby affecting the former adversely. Further, it requires large, self-defeating and unnecessary increases in taxes which may have a dampening influence on investment.

However, the weaknesses of the balanced budget dogma of the classicals led economists to propound the balanced budget theorem. The classical principle of balancing the budget annually is contradictory to the policy of economic stability. For it means that during inflation the government should either increase government expenditure or reduce taxes to balance the budget which would intensify rather than pacify inflation. Since during depression the government revenues decline, the deficit can be eliminated by either increasing taxes or reducing government expenditure. Such a policy would bring the economy to the bottom of the depression. Thus a policy of balanced budgeting would have harmful effect on the economy. In this sense, the balanced budget theorem is superior to the classical doctrine of balanced budgeting.

Some economists, however, favour the Swedish Budget Policy of the 1930s which aims at balancing the budget over the business cycle. Such a policy requires that during inflationary periods the budget should have an excess of tax receipts over expenditure and the same may be utilised for retiring the public debt so that the budget remains a balaned one. On the other hand, during deflationary periods, the budget should have a deficit. Expenditure should be more than the tax receipts and it should be balanced by incurring public debt. Such a policy presupposes a strong government capable of making changes in its expenditure, tax rates and public debt policy. Moreover, it expects of the state to have a machinery

capable of forecasting the cyclical fluctuations accurately. But it is too much to expect of a modern state whose decisions are politically motivated and due to the lack of an accurate machinery to forecast cyclical fluctuations, the balancing of the budget at the appropriate time becomes an impossibility. Economists, therefore, favour compensatory fiscal policy.

EXERCISES

1. Explain the balanced budget multiplier. What are its limitations ?
2. Examine the conditions under which the value of the multiplier is equal to one.
3. Write notes on : Government expenditure multiplier and tax multiplier.
4. Explain how balanced budget multiplier operates when a lumpsum tax and a proportional income tax is levied.

CHAPTER 14

FOREIGN TRADE MULTIPLIER

MEANING

The foreign trade multiplier, also known as the export multiplier, operates like the investment multiplier of Keynes. It may be defined as *the amount by which the national income of a country will be raised by a unit increase in domestic investment on exports.* As exports increase, there is an increase in the income of all persons associated with export industries. These, in turn, create demand for goods. But this is dependent upon their marginal propensity to save (*MPS*) and the marginal propensity to import (*MPM*). The smaller these two marginal propensities are, the larger will be the value of the multiplier, and vice versa.

Its Working

The foreign trade multiplier process can be explained like this. Suppose the exports of the country increase. To begin with, the exporters will sell their products to foreign countries and receive more income. In order to meet the foreign demand, they will engage more factors of production to produce more. This will raise the income of the owners of factors of production. This process will continue and the national income increases by the value of the multiplier. The value of the multiplier depends on the value of *MPS* and *MPM*, there being an inverse relation between the two propensities and the export multiplier.

The foreign trade multiplier can be derived algebraically as follows :

The national income identity in an open economy is

$$Y = C+I+X–M$$

where Y is national income, C is national consumption, I is total investment, X is exports and M is imports.

The above relationship can be solved as :

$$Y - C = I + X - M$$

or $$S = I+X–M \qquad (\because S = Y - C)$$

$$S + M = I + X$$

Thus at equilibrium levels of income the sum of savings and imports (S+M) must equal the sum of investment and export (I+X).

In an open economy the investment component (I) is divided into domestic investment (I_d) and foreign investment (I_f)

$$\because \qquad I=S$$

$$\therefore \qquad I_d+I_f=S \qquad ...(1)$$

Foreign investment (I_f) is the difference between exports and imports of goods and services.

$$I_f=X - M \qquad ...(2)$$

Substituting (2) into (1), we have

$$I_d + X - M = S$$

or $$I_d+X = S+M$$

which is the equilibrium condition of national income in an open economy. The foreign trade multiplier coefficient (K_f) is equal to

$$K_f = \frac{\Delta Y}{\Delta X}$$

and $\Delta X = \Delta S + \Delta M$ $[\because I_a = 0]$

Dividing both sides by ΔY, we get

$$\frac{\Delta X}{\Delta Y} = \frac{\Delta S + \Delta M}{\Delta Y}$$

or $$\frac{\Delta Y}{\Delta X} = \frac{\Delta Y}{\Delta S + \Delta M}$$

or $$K_f = \frac{\Delta Y}{\Delta S + \Delta M} \quad \left(\because K_f = \frac{\Delta Y}{\Delta X}\right)$$

$$K_f = \frac{1}{\frac{\Delta S}{\Delta Y} + \frac{\Delta M}{\Delta Y}} \quad (\because \text{Dividing by } \Delta Y)$$

Let us understand it with the help of an example.

Suppose *MPS*=0.3, *MPM* = 0.2 and Δ*X* (increase in exports) = Rs. 1000 crores, we get

$$K_f = \frac{\Delta Y}{\Delta X} = \frac{1}{MPS + MPM}$$

or $$\Delta Y = \frac{1}{MPS + MPM} \Delta X$$

$$= \frac{1}{0.3 + 0.2} \times 1000 = \text{Rs. } 2000 \text{ crores}$$

It shows that an increase in exports by Rs. 1000 crores has raised national income through the foreign trade multiplier by Rs. 2000 crores, given the values of *MPS* and *MPM*.

Its Assumptions

The foreign trade multiplier is based on the following assumptions :

1. There is full employment in the domestic economy.

2. There is direct link between domestic and foreign country in exporting and importing goods.

3. The country is small with no foreign repercussion effects.

4. It is on a fixed exchange rate system.

5. The multiplier is based on instantaneous process without time lags.

6. There is no accelerator.

7. There are no tariff barriers and exchange controls.

8. Domestic investment (I_d) remains constant.

9. Government expenditure is constant.

10. The analysis is applicable to only two countries.

Diagrammatic Explanation

Given these assumptions, the equilibrium level in the economy is shown in Figure 1, where $S(Y)$ is the saving function and $(S+M)Y$ is the saving plus import function. I_d represents domestic investment and I_d $+X$, domestic investment *plus* exports. $(S+M)Y$ and $I_d + X$ functions determine the equilibrium level of national income OY at point E, where savings equal domestic investment and exports equal imports.

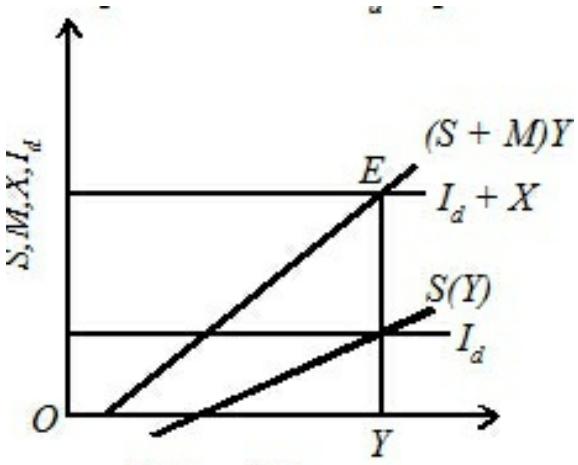


Fig. 1

If there is a shift in the $I_d + X$ function due to an increase in exports, the national income will increase from OY to OY_1, as shown in Figure 2. This increase in income is due to the multiplier effect, *i.e.* $\Delta Y = K_f \Delta X$. The

exports will exceed imports by *sd*, the amount by which savings will exceed domestic investment. The new equilibrium level of income will be OY_1. It is a case of positive foreign investment.

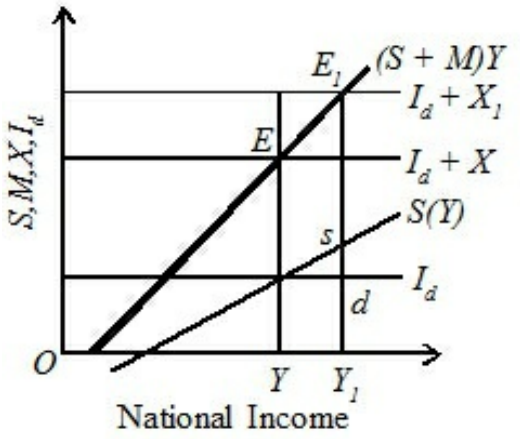


Fig. 2

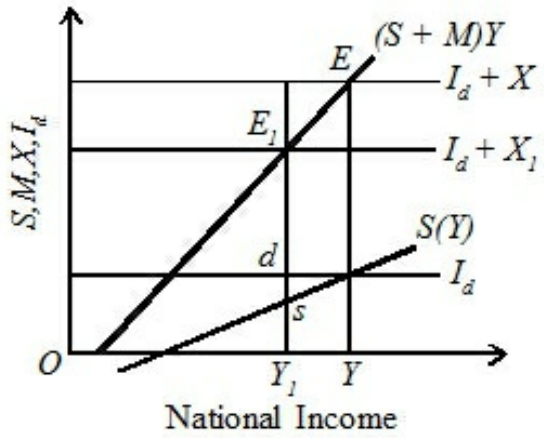


Fig. 3

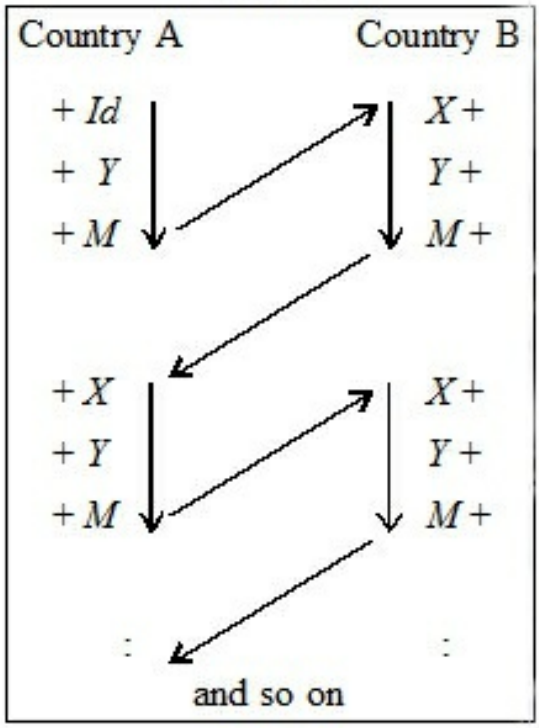
If there is a fall in exports, the export function wil shift downward to I_d+X_1, as shown in Figure 3. In this case imports would exceed exports and domestic investment would exceed savings by *ds*. The level of national income is reduced from *OY* to OY_1. This is the reverse operation of the foreign trade multiplier.

Foreign Repercussion or Backwash Effect

The above analysis of the simple foreign trade multiplier has been studied in the case of one small country. But, in reality, countries are linked to each other indirectly also. A country's exports or imports affect the national income of the other country which, in turn, affects the foreign trade and national income of the first country. This is known as the *Foreign Repercussion or Backwash or Feedback Effect*. The smaller the country is in relation to other trading partner, the negligible is the foreign repercussion. But the foreign repercussion will be high in the case of a large country because a change in the national income of such a country will have significant foreign repercussions or backwash effects. Assuming two large countries *A* and *B* where *A*'s imports are *B*'s exports and vice versa. An increase in *A*'s domestic investment will cause a multiplier increase in its income. This will increase its imports. This increase in *A*'s

imports will be increase in *B*'s exports which will increase income in *B* through *B*'s foreign trade multiplier. Now the increase in *B*'s income will bring an increase in its imports from country *A* which will induce a second round increase in *A*'s income, and so on. This is explained in Table 1. When autonomous domestic investment (I_d) increases in country *A*, its national income increase (+*Y*). It induces country *A* to import more from country *B*. This increases the demand for country *B*'s exports (*X*+). Consequently, the national income in country *B* increases (*Y*+). Now this country imports more (*M*+) from country *A*. As the demand for country *A*'s exports increases (+*X*), its national income (+*Y*) increases further and this country imports more (+*M*) from *B* country. This process will continue in smaller rounds. These are the foreign repercussions or the backwash effects for country A which will peter out and dampen the effects of increase in the original autonomous domestic investment (I_d) in country *A*.

Table 1



The stages of foreign repercussions shown in the above table are explained in Figure 4 Panel I,II and III. In stage I, domestic investment in country *A* increases form I_d to I_{d1} in Panel I. This leads to an upward shift in the $I_d + X$ curve to $I_{d1} + X$. As a result, the new equilibrium point is at E_1 which shows an increase in the national income from *OY* to OY_1. As the national income increases, the demand for imports from country *B* also increases. This means increase in the exports of country *B*. This is shown in Panel II when the $I_d + X$ curve of country *B* shifts upward as $I_d + X_1$. Consequently, the national income in country *B* increases from OY_0 to *OY'* at the higher equilibrium level *E'*. As country *B*'s income increases, its demand for imports from country *A* also increases. This, in turn, leads to

the backwash effect in the form of increase in the demand for exports of country *A*. This is shown in Panel III where the $I_{d1} + X$ curve (of PaneI I) further shifts upwards to $I_{d1} + X_1$ and consequently the national income increases further from OY_1 to OY_2.

This shows how the foreign repercussions in one country affect its own national income and that of the other country which, in turn, again affects in own national income through the backwash effects with greater force.

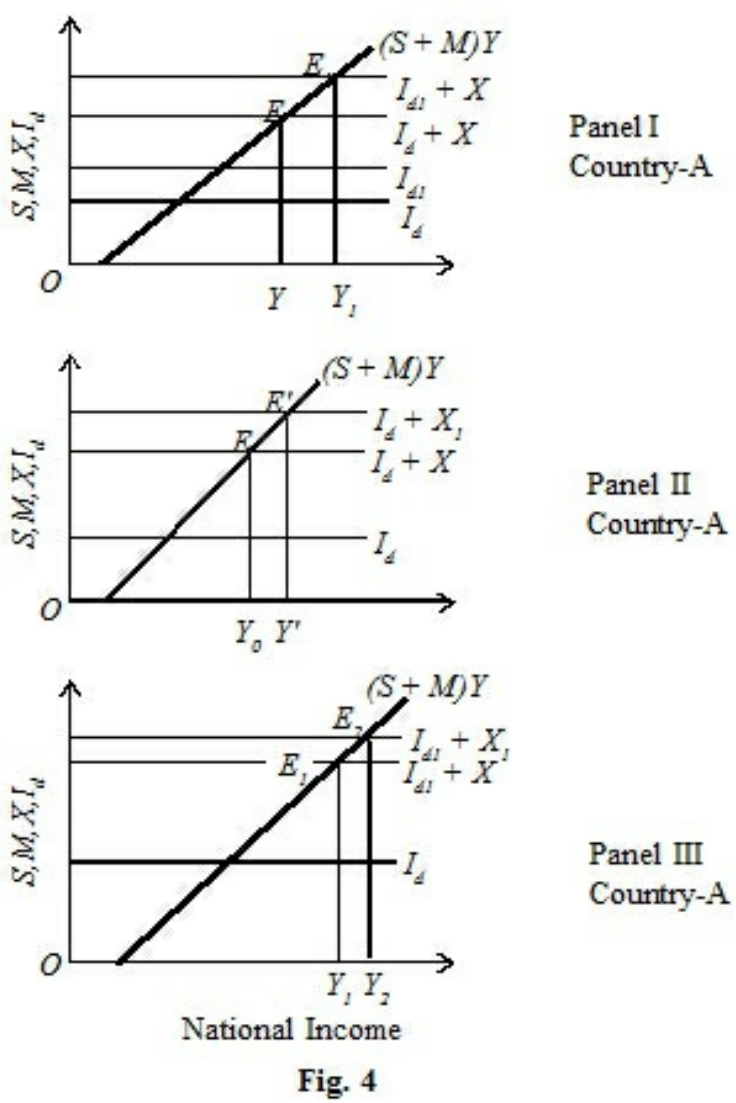


Fig. 4

Implications of Foreign Repercussion

The following are the implications of foreign repercussion effects:

1. The foreign repercussion effects suggest a mechanism for the transmission of income disturbances between trading countries. If a country is small, it will be affected by change in income of other countries, that will alter the demand for its exports. But it will not be able to transmit its own income disturbances to the latter. If a country is large, it may trnasmit its own income disturbances to other countries and, in turn, be affected by income disturbances in them. It implies that a boom or slump in one country has repercussion on the incomes of other countries. Thus swings in business cycles are likely to be internationally contagious, as happened in the 1930s and 2008.

2. The repercussion effects also suggest that since the backwash effects ultimately peter out, automatic income changes cannot eliminate completely the current account BOP deficit or surplus produced by an automatic disturbance.

3.The policy implications of the backwash effects suggest that export promotion policies raise national income in the trading partners at a lower rate than by an increase in domestic investment. The export promotion measures raise national income via the simple foreign trade multiplier, whereas increase in domesic investment policies raise national income many times in multiplier rounds via the repercussion effects.

Criticisms of Foreign Trade Multiplier

The two models of the foreign trade multiplier presented above are based on certain assumptions which make the analysis unrealistic.

1. Exports and Investment not Independent. The analysis of simple foreign trade multiplier is based on the assumption that exports and investment (both domestic and foreign) are independent of changes in the level of national income. But, in reality, this is not so. A rise in exports does not always lead to increase in national income. On the contrary, certain imports, of say capital goods, have the effect of increasing national income.

2. Lagless Analysis. The foreign trade multiplier is assumed to be an instantaneous process whereby it provides the final results. Thus it involves no lags and is unrealistic.

3. Full Employment not Realistic. The analysis is based on the asumption of a fully employed economy. But there is less than full employment in every economy. Thus the foreign trade multiplier does not find clear expression in an economy with less than full employment.

4. Not Applicable to More than two Countries. The whole analysis is applicable to a two-country model. If there are more than two countries, it becomes complicated to analyse and interpret the foreign repercussions of this theory.

5. Neglects Trade Restrictions. The foreign trade multiplier assumes that there are no tariff barriers and exchange controls. In reality, such trade restrictions exist which restrict the operations of the foreign trade multiplier.

6. Neglects Monetary-Fiscal Measures. This analysis is based on the unrealistic assumption that the government expenditure is constant. But governments always interfere through monetary and fiscal policies which affect exports, imports and national income.

Despite these shortcomings, the foreign trade multiplier is a powerful tool of economic analysis which helps in formulating policy measures.

SELECTED READINGS

Eugene Diulio, Macroeconomics, 3/e, 1999.

Richard T. Froyen, Macroeconomics 7/e, 2002.

Roger E.A. Farmer, Macroeconomics, 2/e, 2002.

N. Gregory Mankiw, Macroeconomics, 4/e, 2002.

EXERCISES

1. Explain the Foreign Trade Multiplier without and with foreign repercussions.
2. What is a foreign trade multiplier ? How does it work ?
3. Explain the working and limitations of foreign trade multiplier.
4. Explain the inter-relationship between national income and international trade of a country.

CHAPTER 15

THE PRINCIPLE OF ACCELERATION AND THE SUPER MULTIPLIER

INTRODUCTION

T.N. Carver was the earliest economist who recognised the relationship between changes in consumption and net investment in 1903. But it was Aftalion who analysed this principle in detail in 1909. The term “acceleration principle” itself was first introduced into economics by J. M. Clark[1] in 1917. It was further developed by Hicks, Samuelson, and Harrod in relation to the business cycles.

THE PRINCIPLE OF ACCELERATION

The principle of acceleration is based on the fact that the demand for capital goods is derived from the demand for consumer goods which the former help to produce. The acceleration principle explains the process by which an increase (or decrease) in the demand for consumption goods leads to an increase (or decrease) in investment on capital goods. According to Kurilara, "The accelerator coefficient is the ratio between induced investment and an initial change in consumption expenditure."

1. J.M. Clark, *"Business Acceleration and the Law of Demand, A Technical Factor in Economic Cycles."* J.P.E., 1917. Reprinted in Readings in Business Cycle Theory, 1944.

Symbolically, $v = \Delta I / \Delta C$ or $\Delta I = v\ \Delta C$ where v is the accelerator coefficient, ΔI is net change in investment and ΔC is the net change in

consumption expenditure. If the increase in consumption expenditure of Rs 10 crores leads to an increase in investment of Rs 30 crores, the accelerator coefficient is 3.

This version of the acceleration principle has been more broadly interpreted by Hicks as the ratio of induced investment to changes in output it calls forth. Thus the accelerator v is equal to $\Delta I/\Delta Y$ or the capital-output ratio. It depends on the relevant change in output (ΔY) and the change in investment (ΔI). It shows that the demand for capital goods is not derived from consumer goods alone but from any *direct* demand of national output.

In an economy, the required stock of capital depends on the change in the demand for output. Any change in output will lead to a change in the capital stock. This change equals v times the change in output. Thus $\Delta I = v\Delta Y$, where v is the accelerator. If a machine has a value of Rs 4 crores and produces output worth Rs 1 crore, then the value of v is 4. An entrepreneur who wishes to increase his output by Rs 1 crores every year must invest Rs 4 crores on this machine. This equally applies to an economy where if the value of the accelerator is greater than one, more capital is required per unit of output so that the increase in net investment is greater than the increase in output that causes it. Gross investment in the economy will equal replacement investment *plus* net investment. Assuming replacement investment *(i.e.,* replacement demand for machines due to obsolescence and depreciation) to be constant, gross investment will vary with the level of investment corresponding to each level of output.

The acceleration principle can be expressed in the form of the following equation.[2]

$$I_{gt} = v\,(Y_t - Y_{t-1}) + R$$

$$= v\,\Delta Y_t + R$$

where I_{gt} is gross investment in period t, v is the accelerator, Y_t is the national output in period t, Y_{t-1} is the national output in the previous

period *(t—1)*, and *R* is the replacement investment.

The equation tells that gross investment during period *t* depends on the change in output *(Y)* from period *t* — 1 to period *t multiplied* by the accelerator (*v*) *plus* replacement investment *R.*

In order to arrive at net investment (*In*), *R* must be deducted from both sides of the equation so that net investment in period *t* is

$$I_{nt} = v\,(Y_t - Y_{t-1})$$

$$= v\,\Delta Y_t$$

If $Y_t > Y_{t-1}$ net investment is positive during period *t.* On the other hand, if $Y_t < Y_{t-1}$, net investment is negative or there is disinvestment in period *t.*

---

2. F.S. Brooman, *op. cit.,* p. 169.

Operation of the Acceleration Principle

The working of the acceleration principle is explained in Table I.

Table I: Operation of the Acceleration Principle : v = 4

Period in Years	*Total Output (Y)*	*Required Capital*	*Replacement Investment (R)*		*Net Investment (I_n)*		*Gross Investme (I_g)*
(1)	(2)	(3)	(4)	+	(5)	=	(6)
t	100	400	40		0		40
t+l	100	400	40		0		40
t+2	105	420	40		20		60
t+3	115	460	40		40		80
t+4	130	520	40		60		100
t+5	140	560	40		40		80
t+6	145	580	40		20		60
t+7	140	560	40		—20		20

t+8	130	520	40	—40	0
t+9	125	600	40	—20	20

The table traces changes in total output, capital stock, net investment and gross investment over ten time periods. Assuming the value of the acceleration *v*=4, the required capital stock in each period is 4 times the corresponding output of that period, as shown in column (3). The replacement investment is assumed to be equal to 10 per cent of the capital stock in period *t*, shown as 40 in each time period. Net investment in column (5) equals *v* times the change in output between one period and the preceding period. For example, net investment in period t+3=$v(y_{t+3} - Y_{t+2})$, or 40=4(115—105). It means that given the accelerator of 4, the increase of 10 in the demand for final output leads to an increase of 40 in the demand for capital goods (machines). Accordingly the *total* demand for capital goods (machines) rises to 80 made up of 40 of replacement and 40 of net investment. Thus the table reveals that net investment depends on the change in total output, given the value of the accelerator. So long as the demand for final goods (output) rises net investment is positive. But when it falls net investment is negative. In the table, total output (column 2) increases at an increasing rate from period t+l to $t+4$ and so does net investment (column 5). Then it increases at a diminishing rate from period $t+5$ to $t+6$ and net investment declines from period t+7 to $t+9$, total output falls, and net investment becomes negative.

The acceleration principle is illustrated diagrammatically in Figure 1 where in the upper portion, total output curve *Y* increases at an increasing rate up to $t+4$ period, then at a decreasing rate up to period $t+6$. After this it starts diminishing. The curve $I_{n,}$ in the lower part of the figure shows that the rising output leads to increased net investment upto $t+4$ period because output is increasing at an increasing rate. But when output increases at decreasing rate between $t+4$ and $t+6$ periods, net investment declines. When output starts declining in period $t+7$, net investment becomes negative. The curve I_g represents gross investment of the economy. Its behaviour is similar to the net investment curve. But there is one difference that gross investment is not negative and once it becomes zero in period t+8, the curve I_g again starts rising. This is because despite

net investment being negative, the replacement investment is taking place at a uniform rate.

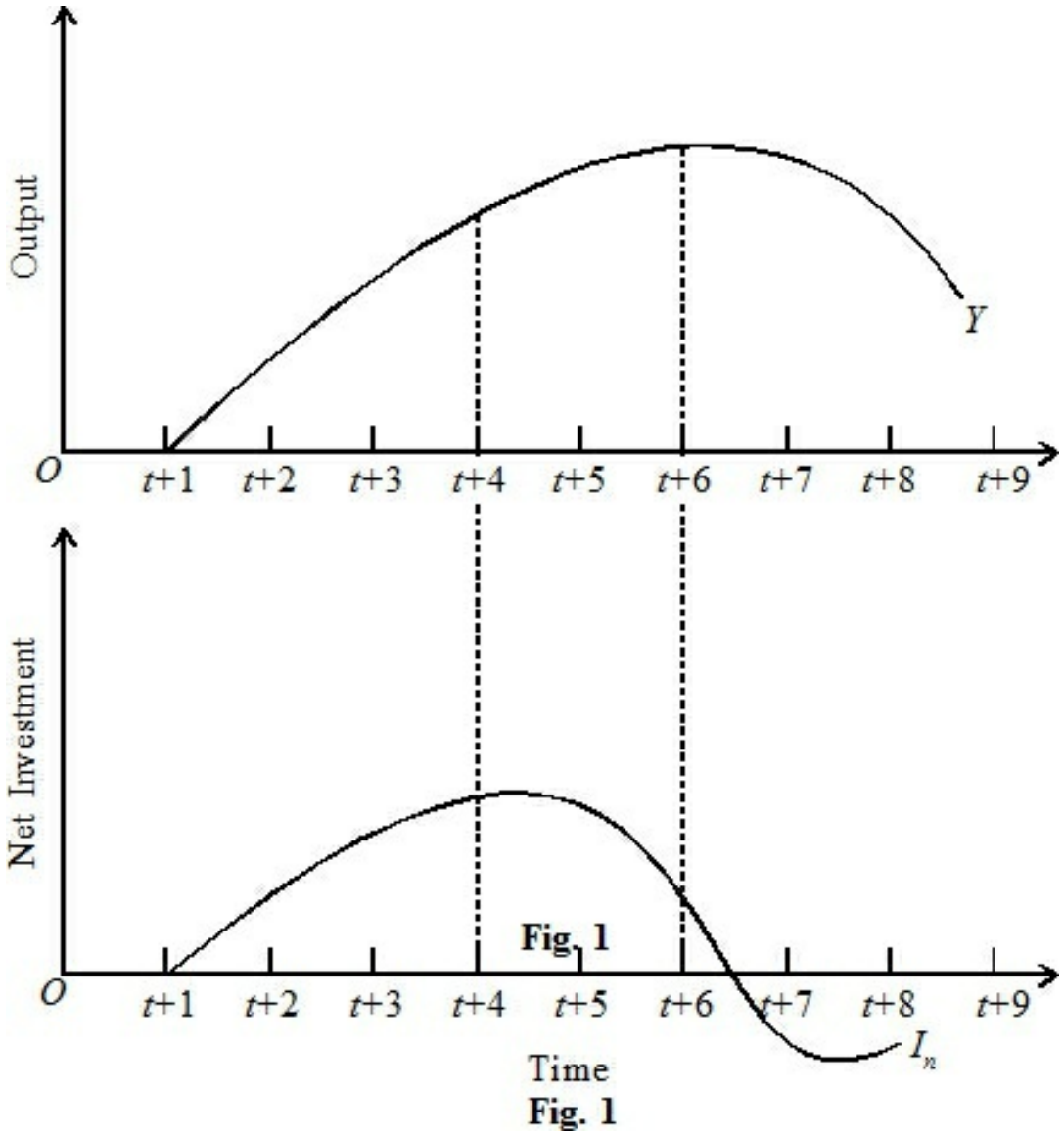


Fig. 1

Assumptions

The acceleration principle is based upon the following assumptions:

1. The acceleration principle assumes a constant capital-output ratio.

2. It assumes that resources are easily available.

3. It assumes that there is no excess or idle capacity in plants.

4. It is assumed that the increased demand is permanent.

5. It also assumes that there is elastic supply of credit and capital.

6. It further assumes that an increase in output immediately leads to a rise in net investment.

Criticisms

The acceleration principle has been criticised by economists for its rigid assumptions which tend to limit its smooth working. The following are its limitations.

1. Capital-Output Ratio not Constant. The acceleration principle is based on a constant capital-output ratio. But this ratio does not remain constant in the modern dynamic world. Inventions and improvements in techniques of production are constantly taking place which lead to increase in output per unit of capital. Or, existing capital equipment may be worked more intensively. Moreover, change in the expectations of businessmen with regard to prices, wages, interest may affect future demand and vary the capital-output ratio. Thus the capital-output ratio does not remain constant but changes in the different phases of the trade cycle.

2. Resources not Elastic. The acceleration principle assumes that the resources should be elastic so that they are employed in the capital goods industries to enable them to expand. This is possible when there is unemployment in the economy. But once the economy reaches the full employment level, the capital goods industries fail to expand due to the non-availability of sufficient resources. This limits the working of the acceleration principle. So this principle will not apply in a recession where excess capacity is found.

3. Idle Capacity in Plants. The acceleration theory assumes that there is no unused (or idle) capacity in plants. But if some machines are not working to their full capacity and are lying idle, then an increase in the demand for consumer goods will not lead to the increased demand for new capital goods. In such a situation the acceleration principle will not work.

4. Difference between Required and Real Capital Stock. It assumes no difference between required and real capital stock. Even if it exists, it ends in one period. But if industries are already producing capital goods at full capacity, it is not possible to end the difference in one period.

5. Does not Explain Timing of Investment. The assumption of the existence of full capacity implies that increased demand for output immediately leads to induced investment. The acceleration principle, therefore, fails to explain the timing of investment. At best it explains the volume of investment. As a matter of fact, there may be a time lag before new investment can be generated. For instance, if the time lag is four years, the effect of new investment will not be felt in one year but in four years.

6. Does not consider Availability and Cost of Capital Goods. The timing of the acquisition of capital goods depends on their availability and cost, and the availability and cost of financing them. The theory does not consider these factors.

7. Acceleration Effect Zero for Installed Equipment. It is assumed that no increase in demand for consumer goods has been foreseen and provided for in previous capital investment. If by anticipating future demand, capital equipment has already been installed, it would not lead to induced investment and the acceleration effect will be zero.

8. Does not Work for Temporary Demand. This theory further assumes that the increased demand is permanent. In case the demand for consumer goods is expected to be temporary, the producers will refrain from investing in new capital goods. Instead they may meet the increased demand by working the existing capital equipment more intensely. So the acceleration will not materialise.

9. Supply of Credit not Elastic. The acceleration principle assumes an elastic supply of credit so that when there is induced investment as a result of induced consumption, cheap credit is easily available for investment in capital goods industries. If cheap credit is not available in sufficient quantities, the rate of interest will be high and investment in capital goods will be very low. Thus the acceleration will not work fully.

10. Neglects Profits as a Source of Internal Funds. This assumption further implies that firms resort to external sources of finance for investment purposes. But empirical evidence has shown that firms prefer internal sources of finance to external sources. The acceleration principle is weak in that it neglects profits as a source of internal finance. As a matter of fact, the level of profits is a major determinant of investment.

11. Neglects the Role of Expectations. The acceleration principle neglects the role of expectations in decision-making on the part of entrepreneurs. The investment decisions are not influenced by demand alone. They are also affected by future anticipations like stock market changes, political developments, international events, economic climate, etc.

12. Neglects the Role of Technological Factors. The acceleration principle is weak in that it neglects the role of technological factors in investment. Technological changes may be either capital-saving or labour-saving. They may, therefore, reduce or increase the volume of investment. Further, as pointed out by Professor Knox, "capital equipment may be bulky and the employment of additional plant is justified only when output has risen considerably. This factor is all the more important because usually what is added is a complex of machines and not a machine."[3]

13. Fails to Explain Lower Turning Point. According to Knox, the acceleration principle is not of much use for explaining lower turning point.

14. Not Precise and Satisfactory. Again, Knox points out that the acceleration, principle is not precise and is unsatisfactory. It is, therefore, inadequate as theory of investment.[4]

Conclusion. Despite these limitations, the principle of acceleration makes the process of income propagation clearer and more realistic than the multiplier theory. The multiplier shows the effect of a change in investment on income via consumption while the acceleration shows the effect of consumption or output on investment and income. Thus the acceleration explains volatile fluctuations in income and employment as a

result of fluctuations in capital goods industries. But it can explain upper turning points better than lower turning points.

THE SUPER-MULTIPLIER OR THE MULTIPLIER-ACCELERATOR INTERACTION

In order to measure the total effect of initial investment on income, Hicks has combined the multiplier and the accelerator mathematically and given it the name of the *super-multiplier*.[4] The combined effect of the multiplier and the accelerator is also called the *leverage effect* which may lead the economy to very high or low level of income propagation.

The super-multiplier is worked out by combining both induced consumption (cY or $\Delta C/\ \Delta Y$ or MPC) and induced investment (vY or $\Delta I/\ \Delta Y$ or MPI). Hicks divides the investment component into autonomous investment and induced investment so that investment $1 = I_a + vY$, where I_a is autonomous investment and vY is induced investment.

---

3. A.D. Knox, 'the Acceleration Principle and the Theory of Investment: A Survey'. Reprinted in M.G. Mueller, *Readings in Macroeconomics*, 2/e, 1971.

4. J.R. Hicks, *A Contribution on the Theory of the Trade Cycle*, 1950, p. 62.

Since $Y = C + I$

Therefore, $\Delta Y = c\,\Delta\,Y + \Delta Ia + v\Delta Y$

$$\Delta Y - c\Delta Y - v\Delta Y = \Delta Ia$$

$$\Delta Y\,(1-c-v) = \Delta Ia$$

$$\frac{\Delta Y}{\Delta I_a} = \frac{1}{1-c-v} = \frac{1}{s-v}$$

or $$K_S = \frac{1}{1-c-v} = \frac{1}{s-v}$$

where K_S is the super-multiplier, c is the marginal propensity to consume, v the marginal propensity to invest, and s is the marginal propensity to save ($s = 1 - c$).

The super-multiplier tells us that if there is an initial increase in autonomous investment, income will increase by K_S times the autonomous investment. So the super-multiplier in general form will be

$$\Delta Y = \frac{1}{1-c-v}\Delta Ia$$
$$= \mathrm{K_S}\,\Delta Ia$$

Let us explain the combined operation of the multiplier and the accelerator in terms of the above equation. Suppose c = 0.5, v = 0.4 and autonomous investment increases by Rs. 100 crores. The increase in aggregate income will be

$$\Delta Y = \frac{1}{1-0.5-0.4}\times 100$$

$$= \frac{1}{0.1}\times 100 = 10\times 100 = 1000$$

It shows that a rise in autonomous investment by Rs 100 crores has raised income to Rs. 1000 crores. The simple multiplier would have raised income to only Rs. 200 crores, given the value of K the multiplier as 2 (since MPC = 0.5). But the multiplier combined with the accelerator (K_S = 10) has raised income to Rs. 1000 crores which is higher than generated by the simple multiplier. Table II explains how the process of income propagation via the multiplier and the accelerator with the value of the super-multiplier K_S = 10 leads to a rise in income to Rs. 1000 crores with an initial investment of Rs. 100 crores.

In period t+1 constant investment of 100 is injected into the economy but there is no immediate induced consumption or investment. In period t+2, induced consumption of 50 takes place out of the income 100 of period t+1, since the marginal propensity to consume is 0.5, while there is an induced investment of 40 out of 100 income (v being 0.4). The increase in income from period 1 to 2 is (50+40) = 90. The increase in income in different periods can be calculated as $\Delta Y_{t+2} = c\,\Delta Y_{t+1} + v\,\Delta Y_{t+1} = 0.5\times100 + 0.4\times100 = 90$. Similarly, the increase in income in period t+3 can be calculated as $\Delta Y_{t+3} = c\Delta Y_{t+2} + v\,\Delta Y_{t+2} = 0.5\times90+0.4\times90=45+36=81$. The

total increase in income (column 6) is arrived at by adding the increase in income (column 5) of the current period to the total increase in income (column 6) of the previous period. For instance, the total increase in income (column 6) in period t + 2 of 190 is arrived at by adding the increase in income (column 5) of this period to the total increase in income 100 (of column 6) of the previous period t+1. Similarly, the total increase in income in period t+3 of 271 = increase in income of 81 in this period *plus* 190 of column 6 of period t+2. This cumulative process of income propagation continues till in period t+n, induced consumption, induced investment and increase in income dwindle to zero. If we add up the increase in consumption, investment and income from period t+1 to t+n, the total income increases to Rs 1000 crores, total consumption to Rs 500 crores and total investment to Rs 400 crores, given the initial investment of Rs 100 crores.

Table II : Multiplier-Accelerator Interaction

(Rs Cror

Period (t)	Initial Investment	Induced Comsumption (c=0.5)	Induced Investment (v=0.4)	Increase in Income (ΔY=c+v)	Total Increas in Incon
(1)	(2)	(3)	(4)	(5)	(6)
t+0	0	0	0	0	0
t+1	100	—	—	100	100
t+2	100	50	40	90	190
t+3	100	45	36	81	271
t+4	100	40.5	32.4	72.9	343.9
t+5	100	36.45	29.16	65.61	409.51
....	....	....	...	...	...
t+n	100	0	0	0	1000

The dynamic path of income is shown in the adjoining Fig. 2. Income is measured vertically and time horizontally. The curve OY_t shows the time path of income with a super-multiplier of 10. The curve rises with time and reaches the new equilibrium level of income Y_1 and flattens out. It indicates that income increases at a decreasing rate.

USE OF MULTIPLIER-ACCELERATOR INTERACTION IN BUSINESS CYCLES

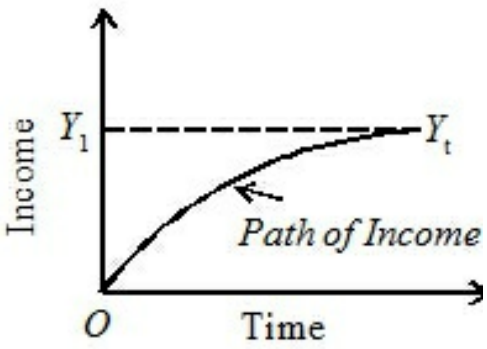


Fig. 2

However, with different values of MPC and the accelerator, the multiplier-accelerator may show different results in terms of cyclical fluctuations. Suppose the *MPC* is 0.5 and the accelerator coefficient is 2. Given the same assumptions and the initial investment of Rs 100 crores, let us study how changes in income take place. Table III explains this process of income propagation.

Table III : Multiplier-Accelerator Interaction

(Rs Cror

Time (t)	Initial Investment	Induced Consumption (c=0.5)	Induced Investment (v=2)	Increase in Income (columns 2+3+4
(1)	(2)	(3)	(4)	(5)
0	0	0	0	0
t+1	100	—	—	100
t+2	100	50	100	250
t+3	100	125	150	375
t+4	100	187.50	125	412.50
t+5	100	206.25	37.50	343.75
t+6	100	171.88	–68.74	203.14
t+7	100	101.57	–140.62	60.95
t+8	100	30.48	–142.18	–11.70
t+9	100	–5.48	–72.66	21.49
t+10	100	10.75	33.20	–143.95

Source : E. Schenneider, Money, Income and Employment, 1968

Table III reveals that in period *t*+1 there is an increase of Rs. 100 crores by the amount of initial investment. This increase in income leads to a rise in consumption of Rs 50 crores (column 3) in period *t*+2 because the value

of *MPC* is 0.5. This rise in consumption induces investment of Rs 100 crores = 50×2 (column 4), the accelerator coefficient being 2. And income increases to Rs 250 crores (column 2+column 3+column 4). This increased income, in turn, leads to an increase in consumption of Rs 125 crores in *t*+3 period, being one-half of Rs 250 crores as the *MPC* is 0.5. But consumption in period *t* is a function of income of the previous period. Therefore, the actual increase in consumption in period *t*+3 and *t*+2 i.e. 125–50=75. If we multiply this increase in consumption 75 by the value of the accelerator 2, we get induced investment of 150=75×2 (column 4) in period *t*+3. Thus the total of columns 2+3+4 gives increase in income of Rs 375 crores in period *t*+3. This increased income leads to induced consumption of 187.50 (column 3) in period *t*+4, since *MPC*=0.5. The difference of induced consumption of period *t*+4 and *t*+3 (187.50 minus 125) is 62.50 which multiplied by the value of the accelerator 2 gives the figure of 125 of induced investment (column 4). And the total of columns 2, 3 and 4 gives the increase in income of Rs 412.50 crores (column5) in period *t*+4, and so on. The increase in income is the highest in period *t*+4 which shows the peak of the cycle. Thereafter, it starts falling till it reaches the bottom or trough when income is minus Rs 11.70 crores in period *t*+8. From period *t*+9, it again starts rising which shows the revival phase of the cycle. This behaviour of income as a result of the combined operation of the multiplier and the accelerator reveals that income first rises, then falls and again rises at constant amplitudes. The actual behaviour of the cycle, however, depends on the values of the multiplier and the accelerator, as shown by Samuelson in his model.[5]

Kurihara[6] points out that a less than unity marginal propensity to consume provides an answer to the question. Why does the cumulative process come to a stop *before* a complete collapse or before full employment? According to Hansen, this is due to the fact that a large part of the increase in income in each period is not spent on consumption in each successive period. This eventually leads to a decline in the volume of induced investment and when such a decline exceeds the increase in induced consumption, a decline in income sets in. Thus, writes Hansen, "It is the marginal propensity to save which calls a halt to the expansion process even when the expansion is intensified by the process of acceleration on top of the multiplier process."[7]

SELECTED READINGS

M.G. Mueller (ed.), *Readings in Economics,* 1966, articles Ch. 9 by A.D Knox, and Ch. 18 by P.A. Samuelson.

Gardner Ackley, *Macroeconomic Analysis,* 1961, Ch. 17.

F.S. Brooman, *Macroeconomics,* 1970, Ch. VII.

T.F.Dernberg and D.M. McDougall, *Macroeconomics,* 5/e, 1976, Ch. 14.

EXERCISES

1. Explain the concept of Super-Multiplier. Describe its working in a free, open economy.
2. Describe the principle of acceleration. What are its limilations?
3. Discuss the assumptions necessary for the smooth working of the acceleration principle.
4. Explain the multiplier-accelerator interact con and give its uses in business cycles.
5. Explain the interaction between the multiplier and the accelerator in the theory of income determination. How far is it possible to provide a satisfactory explanation of the trade cycle on its basis ?

---

5. For figure relating to the above values of c and v, students should refer to Figure 1 (C) in Samuelson's Model of Trade Cycle in the chapter on Business Cycles. Figures relating to the other values of c and v are also given in the same chapter.

6. K.K. Kurihara, *Monetary and Public Policy*, pp.233-34.

CHAPTER 16

SOME NEW THEORIES OF INVESTMENT

INTRODUCTION

In the earlier chapters, we discussed three theories of investment. First, the classical theory that investment is a function of the rate of interest. Second, the Keynesian theory that investment is a function of income and rate of interest. Third, the acceleration theory that investment is a function of the change in output or income.

In this chapter, we shall discuss some of the post-Keynesian theories of investment and refinements in the accelerator theory. First, we explain the simple accelerator principle in its crudest form which has already been discussed in the previous chapter in detail. It has come to be known as the *naive accelerator*.

THE ACCELERATOR THEORY OF INVESTMENT

The accelerator principle states that an increase in the rate of output of a firm will require a proportionate increase in its capital stock. The capital stock refers to the desired or optimum capital stock, K^*. Assuming that capital-output ratio is some fixed constant, v, the optimum capital stock is a constant proportion of output so that in any period t,

$$K_t^* = vY_t$$

where K_t^* is the optimal capital stock in period t, v (the acceleator) is a positive constant, and Y_t is output in period t.

Any change in ouput will lead to a change in the capital stock. Thus

$$K_t^* - K_{t-1}^* = v(Y_t - Y_{t-1})$$

and

$$I_{nt} = v\ (Y_t - Y_{t-1}) \qquad \left[\because I_{nt} = K_t^* - K_{t-1}^*\right]$$
$$= v\ \Delta Y_t$$

where $\Delta Y_t = Y_t - Y_{t-1}$ and I_{nt} is net investment.

This equation represents the *naive accelerator.*

In the above equation, the level of net investment is proportional to change in output. If the level of output remains constant ($\Delta Y = 0$), net investment would be zero. For net investment to be a positive constant, output must increase. This is illustrated in Figure 1 where in the upper portion, the total output curve Y increases at an increasing rate upto $t + 4$ period, then at a decreasing rate upto period $t + 6$. After this, it starts diminishing. The curve I_n in the lower part of the figure, shows that the rising output leads to increased net investment upto t+4 period because output is increasing at an increasing rate. But when ouput increases at a decreasing rate between t+4 and t+6 periods, net investment declines. When output starts declining in period t+7, net investment becomes negative. The above explanation is based on the assumption that there is symmetrical reaction for increases and decreases of output.

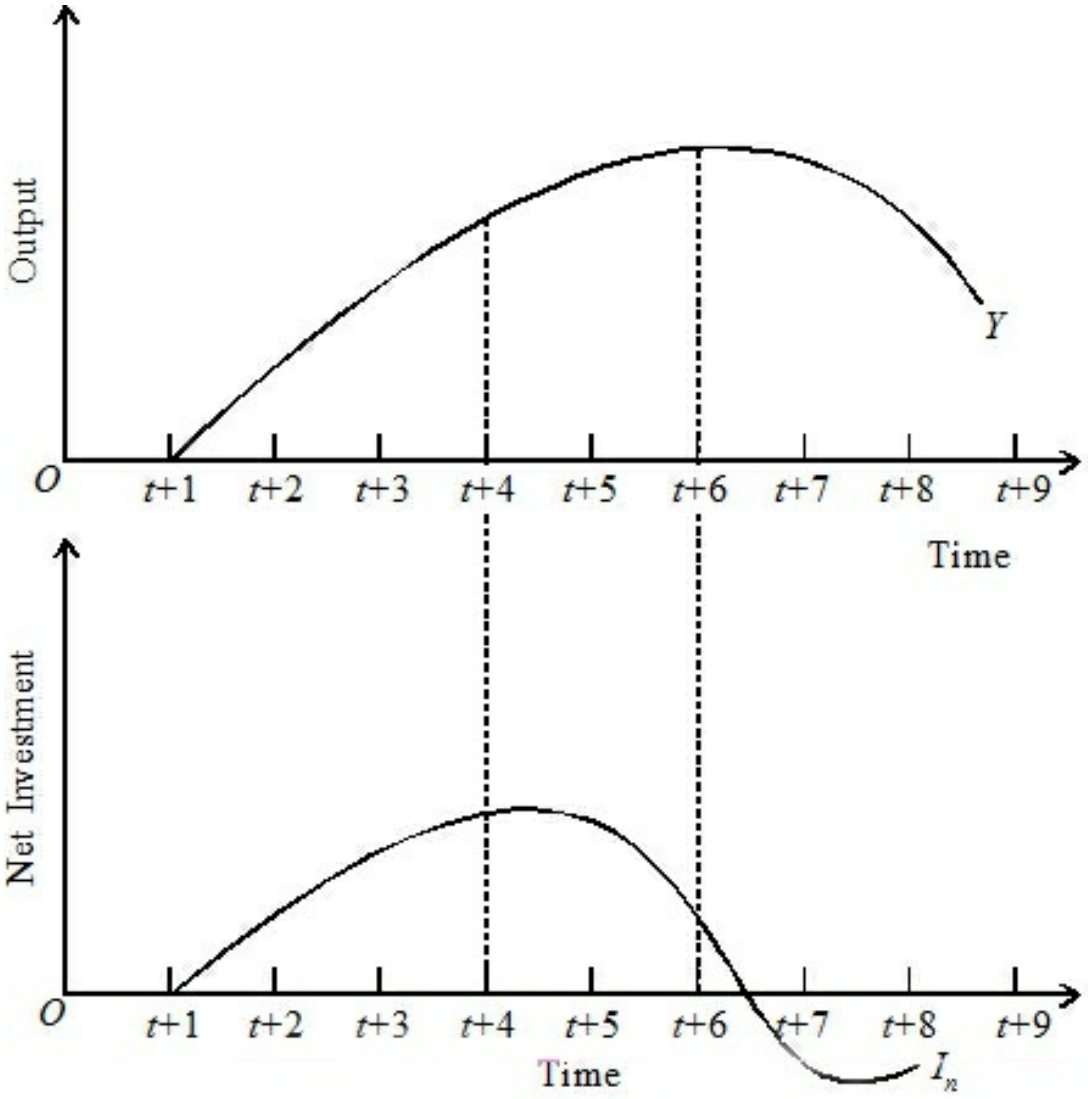


Fig. 1.

In the simple acceleration principle, the proportionality of the optimum capital stock to output is based on the assumption of fixed technical coefficients of production. This is illustrated in Figure 2 where Y and Y_1 are the two isoquants. The firm produces Y output with K^* optimal capital stock. If it wants to produce Y_1 output, it must increase its optimal capital stock to K^*_t . The ray OR shows constant returns to scale. It follows that if the firm wants to double its output, it must increase its optimal capital stock by two-fold.

Eckaus[1] has shown that under the assumption of constant returns to scale, if the factor-price ratios remain constant, the simple accelerator would be

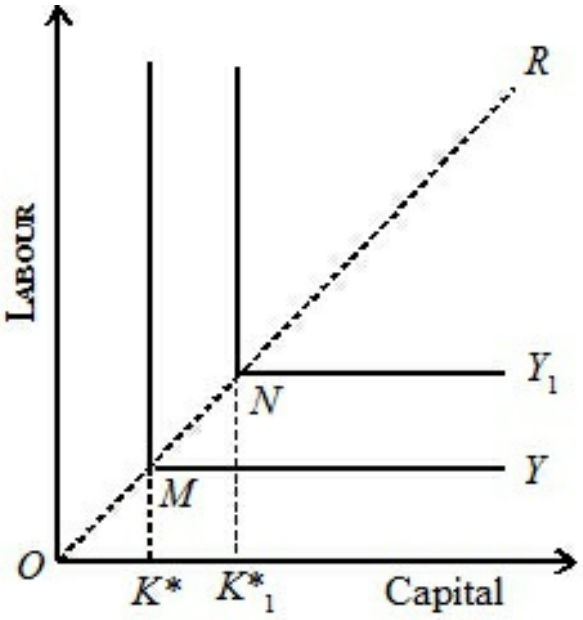


Fig. 2.

The line *OR* joining the points of tangency *e*, e_1, and e_2 is the firms's expansion path which shows investment to be proportional to the change in output when capital is optimally adjusted between the iosquants and isocosts.*

THE FLEXIBLE ACCELERATOR THEORY OR LAGS IN INVESTMENT

The flexible accelerator theory removes one of the major weaknesses of the simple acceleration principle that the captal stock is optimally adjusted without any time lag. In the flexible accelerator, there are lags in the adjustment process between the level of output and the level of capital stock. This theory is also known as *the capital stock adjustment model.* The theory of flexible accelerator has been developed in various forms by Chenery[2], Goodwin,[3] Koyck[4] and Junankar. But the most accepted approach is by Koyck.

constant. Suppose the firm's production involves the use of only two factors, capital and labour whose factor-price ratios are constant. In Figure 3, *Y*, Y_1 and Y_2 are the firms's isoquants and *C*, C_1 and C_2 are the isocost lines which are parallel to each other, thereby showing constant costs. If the firm decides to increase its ouput from *Y* to Y_1, it will have to increase the units of labour from *L* to L_1 and of capital from K^* to K_1^* and so on.

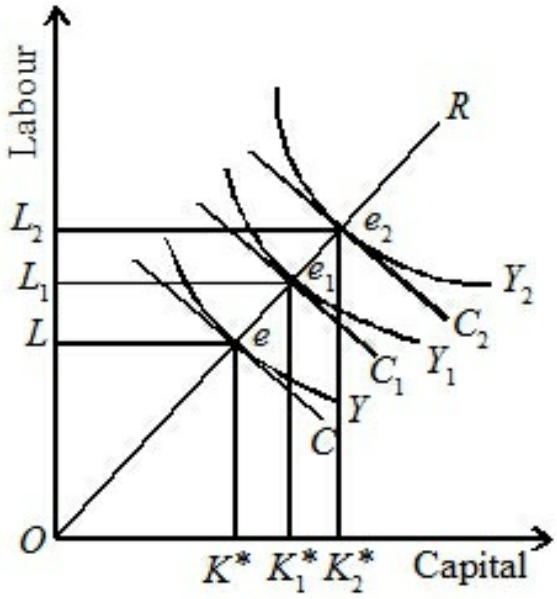


Fig. 3

Junankar[5] has discussed the lags in the adjustment between output and capital stock. He explains them at the firm level and extends them to the aggregate level. Suppose there is an increase in the demand for output. To meet it, first the firm will use its inventories and then utilise its capital stock more intensively. If the increase in the demand for output is large and persists for some time, the firm would increase its demand for capital stock. This is the *decision-making lag*. There may be the *administrative lag* of ordering the capital. As capital is not easily available and in abundance in the financial capital market, there is the *financial lag* in raising finance to buy capital. Finally, there is the *delivery lag* between the ordering of capital and its delivery. Assuming "that different firms have different decision and delivery lags then in aggregate the effect of an increase in demand on the capital stock is distributed over time. This implies that the capital stock at time *t* is dependent on all the previous levels of output, i.e.

---

* *For Assumptions and Criticisms of the* Accelerator Theory of Investment *readers should refer to the previous chapter.*

2. *H.B. Chenery, "Overcapacity and the Acceleration Principle",* Econometrica, *Jan. 1952.*

3. *R.M. Goodwin, "The Nonlinear Accelerator and Persistence of Business cycles",* Econometrica *January 1951.*

4. *L.M. Koyck,* Distributed Lags and Investment Analysis, *1954.*

5. *P.N. Junankar,* Investment : Theories and Investment, *1972.*

$$K_t = f(Y_t, Y_{t-1}.....,Y_{t-n}).$$

This is illustrated in Figure 4 where initially in period t_0, there is a fixed relation between the capital stock and the level of output. When the demand for output increases, the capital stock increases gradually after the decision and delivery lags, as shown by the *K* curve, depending on the previous levels of output. The increase in output is shown by the curve *Y*. The dotted line *K** is the optimal capital stock which equals the actual capital stock *K* in period *t*.

Koyck's Approach

Koyck's approach to the flexible accelerator assumes that the actual capital stock depends on all past output levels with weights declining geometrically. Accordingly,

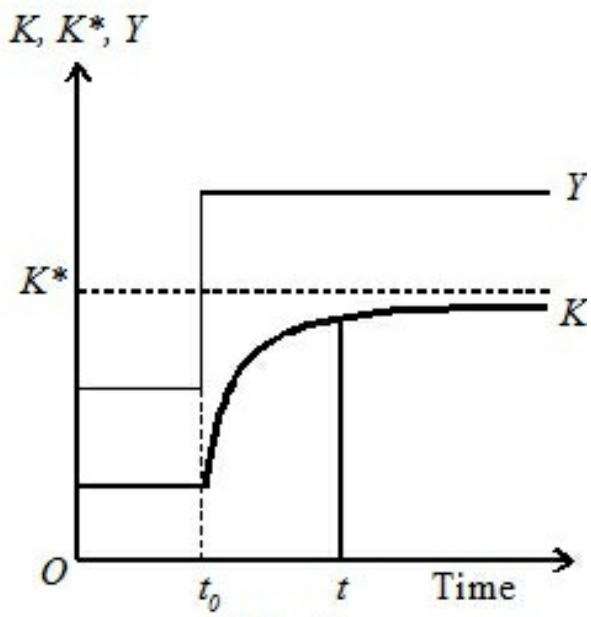


Fig. 4

$$K_t = v\,(1-\lambda)\,(Y_t + \lambda Y_{t-1} + \lambda^2 Y_{t-2} + \ldots\ldots \lambda^n Y_{t-n}) \qquad \ldots\ldots(1)$$

where $0<\lambda<1$. If there is no change in income and it is equal to $\bar{Y}$ the expected volume of output also remains unchanged, then

$$\bar{K} = v\,(1-\lambda)\,(\bar{Y} + \lambda \bar{Y} + \lambda^2 \bar{Y} + \ldots\ldots \lambda^n \bar{Y})$$
$$= v\,(1-\lambda)\,\bar{Y}\,(1 + \lambda + \lambda^2 + \ldots\ldots + \lambda^n) \qquad \ldots\ldots(2)$$

where $(1+\lambda+\lambda^2+\ldots\ldots+\lambda^n) = 1/(1-\lambda)$ are the weights in geometric series and equation (2) becomes

$$\bar{K} = v\bar{Y}\,(1-\lambda) \times 1/(1-\lambda)$$

or $$\bar{K} = v\,\bar{Y}$$

If equation (1) is valid, then K_{t-1} is also true. Therefore, we can rewrite equation (1) as

$$K_{t-1} = v\,(1-\lambda)\,(Y_{t-1} + \lambda^2 Y_{t-2} + \lambda^3 Y_{t-3} + \ldots\ldots + \lambda^n Y_{t-n-1}).$$

Multiplying by λ we have

$$\lambda K_{t-1} = v\,(1-\lambda)\,(\lambda Y_{t-1} + \lambda^2 Y_{t-2} + \lambda^3 Y_{t-3} + \ldots\ldots + \lambda^{n+1} Y_{t-n-1}) \qquad \ldots(3)$$

Subtracting equation (3) from equation (1), we get

$$K_t - \lambda K_{t-1} = v\,(1-\lambda)\,(Y_t + \lambda^{n+1} Y_{t-n-1}).$$

Since the term λ^{n+1} tends to zero, the above equation becomes

$$K_t - \lambda\, K_{t-1} = (1-\lambda)\,vY_t$$

or $$K_t = (1-\lambda)\,vY_t + \lambda K_{t-1} \qquad \ldots(4)$$

This process of rewriting equation (1) as equation (4) is called the *Koyck transformation.*

Net investment is the change in the stock of capital, $K_t - K_{t-1}$. Therefore, subtract K_{t-1} from both sides of the equation to get the expression net investment,

$$K_t - K_{t-1} = (1 - \lambda)\, vY_t + \lambda K_{t-1} - K_{t-1}$$
$$I_{nt} = (1 - \lambda)\, vY_t + K_{t-1}(\lambda - 1)$$
or $$I_{nt} = (1 - \lambda)\, vY_t - (1 - \lambda)\, K_{t-1} \quad \text{.....(5)}$$

The net investment $(K_t - K_{t-1})$ is called the *distributed lag accelerator* which is inversely related to the capital stock of the previous period and is positively related to the output level.

To convert net investment to gross investment, add depreciation (D_t) to both sides of equation (5),

$$I_{nt} + D_t = I_{gt} = (1 - \lambda)\, vY_t - (1 - \lambda)\, K_{t-1} + D_t \quad \text{...(6)}$$

Depreciation is assumed to be proportional to last year's capital stock and is estimated by $D_t = \delta K_{t-1}$. By adding this to equation (6), gross investment (I_{gt}) is

$$I_{gt} = (1 - \lambda)\, vY_t - (1 - \lambda)\, K_{t-1} + \delta K_{t-1}$$
$$= (1 - \lambda)\, vY_t - [(1 - \lambda) + \delta] K_{t-1}$$
$$= (1 - \lambda)\, vY_t - (1 - \lambda\delta)\, K_{t-1} \quad \text{...(7)}$$

This equation represents the *flexible accelerator* or the *stock adjustment principle*. This suggests that "net investment is some fraction of the difference between *planned* capital stock and *actual* capital stock in the previous period...The coefficient $(1 - \lambda)$ tells us how rapidly the adjustment takes place. If $\lambda = 0$ [i.e. $(1 - \lambda) = 1$] then adjustment takes place in the unit period".

To conclude, the flexible accelerator is a very important contribution to the theory of investment which solves the problem of lags in investment demand. It not only incorporates the effects of lags but also of depreciation and excess capacity in the capital stock adjustment.

Its Comparison with Naive Accelerator

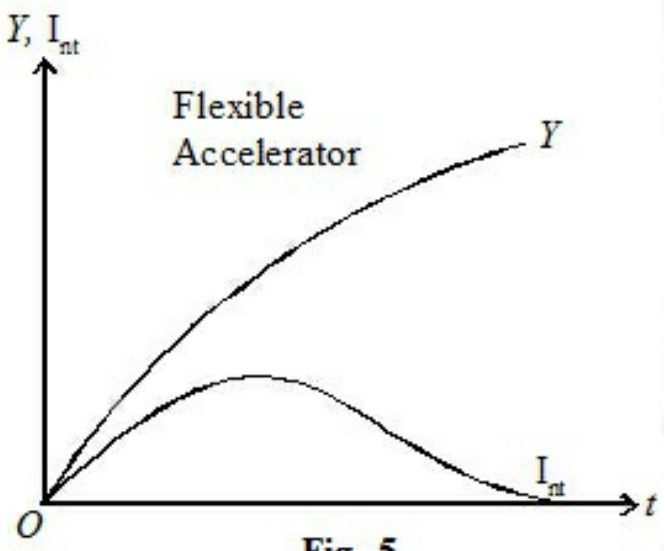


Fig. 5.

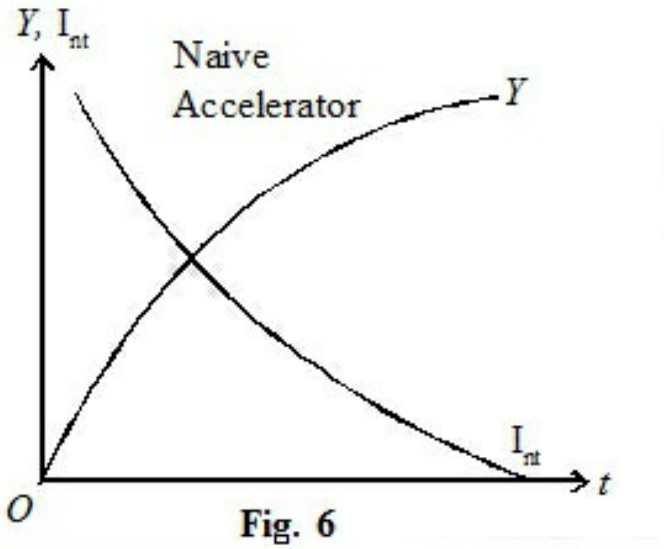


Fig. 6

Since the flexible accelerator and naive accelerator are both accelerators, their long-run response of investment to a change in output will be similar. Let us consider a situation where output (Y) is rising at a decreasing rate and ultimately stops rising at a high level. In the case of the flexible accelerator, net investment will increase during several periods before the negative effect of the increased capital stock outweighs the positive effect of further increases in output and ultimately net investment will become zero. This is shown in Figure 5. On the other hand, in the case of the naive accelerator, net investment will be decreasing continuously and will also become zero, as shown in Figure 6. In both the accelerators, gross investment will be equal to depreciation.

THE PROFITS THEORY OF INVESTMENT

The profits theory regards profits, in particular undistributed profits, as a source of internal funds for financing investment. Investment depends on profits and profits, in turn, depend on income. In this theory, profits relate to the level of current profits and of the recent past. If total income and total profits are high, the retained earnings of firms are also high, and vice versa, Retained earnings are of great inportance for small and large firms

when the capital market is imperfect because it is cheaper to use them. Thus if profits are high, the retained earnings are also high. The cost of capital is low and the optimal capital stock is large. That is why firms prefer to reinvest their extra profit for making investments instead of keeping them in banks in order to buy securities or to give dividends to shareholders. Contrariwise, when their profits fall, they cut their investment projects. This is the *liquidity version* of the profits theory.

Another version is that the optimal capital stock is a function of expected profits. If the aggregate profits in the economy and business profits are rising, they may lead to the expectation of their continued increase in the future. Thus expected profits are some function of actual profits in the past,

$$K_t^* = f(\pi_{t-1})$$

where K_t^* is the optimal capital stock and $f(\pi_{t-1})$ is some function of past actual profits.

Edward Shapiro[6] has developed the profits theory of investment in which total profits vary directly with the income level. For each level of profits, there is an optimal capital stock. The optimal capital stock varies directly with the level of profits. The interest rate and the level of profits, in turn, determine the optimal capital stock. For any particular level of profits, the higher the interest rate, the smaller will be the optimal capital stock, and vice versa. This version of the profits theory is explained in terms of Figure7. The curve *Z* in Panel (A) shows that total profits vary directly with income. When the income is Y_1, profits are P_1 and with increase in income to Y_2, profits rise to P_2**.** Panel (B) shows that the interest rate and the profits level determine the capital stock. At P_2 profits levels and *r*6% interest rate, the actual capital stock is K_2 and at the lower profits level P_1 and interest rate *r*6%, the actual capital stock declines to K_1.

In Panel (C), the *MEC* curve is drawn for each level of profits, given the actual capital stock and the rate of interest. As such, the curve MEC_1 relates the profits level P_1 to the optimal capital stock K_t^* when r6% is the

interest rate. The higher curve MEC_2 relates the profit level P_2 to the higher optimal capital stock K_2*, given the same rate of interest r 6%.

Suppose that the level of profits is P_1, the market interest rate is r6% and the actual capital stock is K_1. With this combination of the variables, the optimal capital stock in Panel (C) is K_1* so that the actual capital stock, $K_1 = K_1$*, the optimal capital stock. As a result, net investment is zero. But there is still I_1 replacement investment at r6%, as indicated by MEI_1 curve in Panel (D). The combination of I_2 investment and Y_1 income level establishes point A on the investment curve I in Panel (E) of the figure.

---

6. *Edward Shapiro*, op.cit., *Ch. 18.*

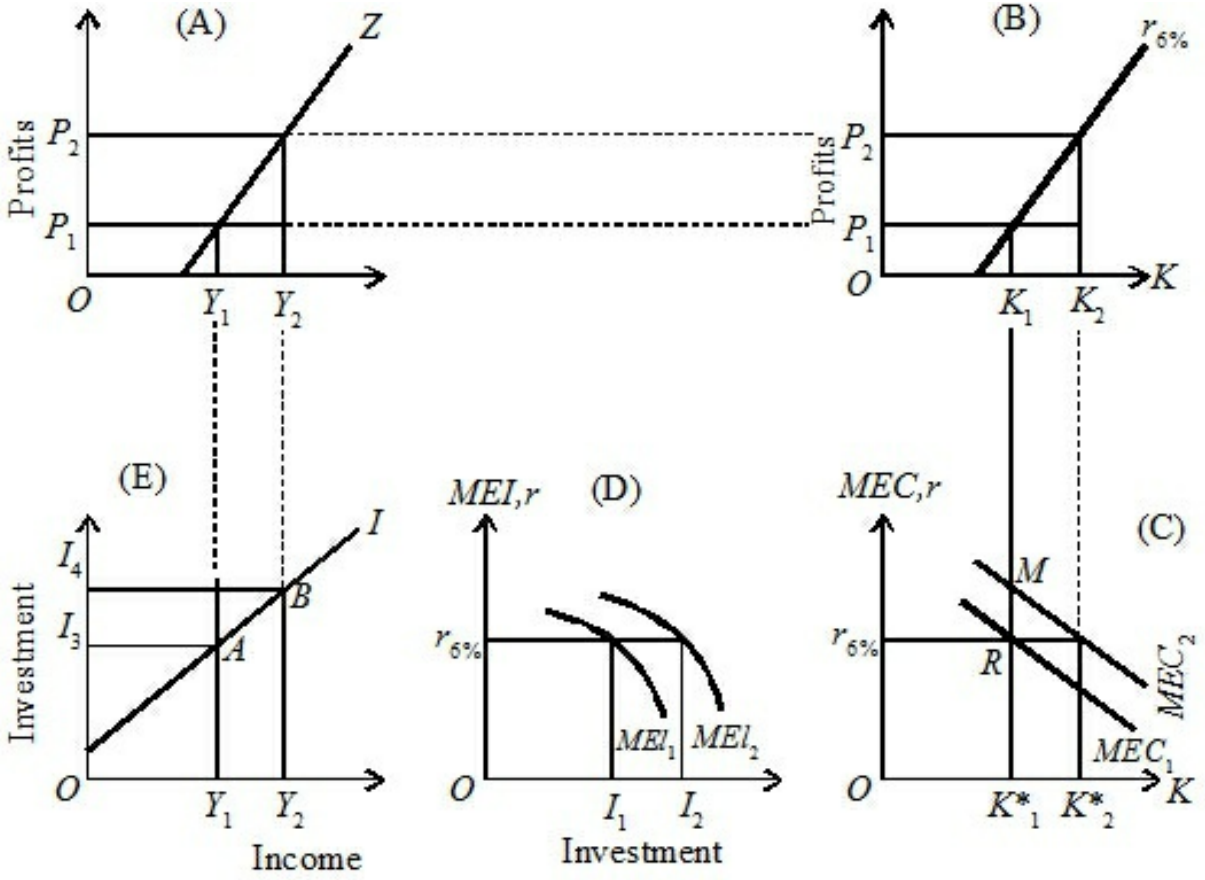


Fig. 7

Now begin with P_2 level of profits and Y_2 income level in Panel (A) so

that at *r*6% interest rate in Panel (C), the optimal capital stock is K_2*. Assuming again that the actual capital stock is K_t, the optimal capital stock is greater than the actual, K_2*>K_1 at this profit-income combination. Here the MEC_2 is higher than *r*6% interest rate by *RM*. As a result, the MEI_1 curve shifts upward to MEI_2 in Panel (D). Since K_2*>K_1, net investment is positive. This is shown by I_1–I_2 in Panel (D). So when profits increase to P_2 with the rise in income to Y_2, the optimal capital stock K_2* being greater than the actual capital stock K_1, at *r*6% interest rate, investment increases from I_3 to I_4 in Panel (E) which is equal to net investment I_1I_2 in Panel (D). The combination of I_4 and Y_2 establishes point *B* on the upward sloping *I* curve.

To sum up, in the profits theory of investment, the level of aggregate profits varies with the level of national income, and the optimal capital stock varies with the level of aggregate profits. If at a particular level of profits, the optimal capital stock exceeds the actual capital stock, there is increase in investment to meet the demand for capital. But the relationships between investment and profits, and between aggregate profits and income are not proportional.

Its Criticism

The theory is based on the assumption that profits are related to the level of current profits and of the recent past. But there is no possibility that the firm's current profit of this year or of the next few years can measure the profits of the next year or of the next few years. A rise in current profits may be the result of unexpected changes of a temporary nature. Such temporary profits do not induce investment.

DUESENBERRY'S ACCELERATOR THEORY OF INVESTMENT

J.S. Duesenberry in his book *Business Cycles and Economic Growth* presents an extension of the simple accelerator and integrates the profits theory and the acceleration theory of investment. Duesenberry has based his theory on the following propositions: (1) Gross investment starts

exceeding depreciation when capital stock grows. (2) Investment exceeds savings when income grows. (3) The growth rate of income and the growth rate of capital stock are determined entirely by the ratio of capital stock to income. He regards investment as a function of income (Y), capital stock (K), profits (π) and capital consumption allowances (R). All these are independent variables and can be represented as

$$I = f\,(Y_{t-1}, K_{t-1}, \pi_{t-1}, R_t\,)$$

where t refers to the current period and (t–1) to the previous period. According to Duesenberry, profits depend positively on national income and negatively on capital stock.

$$\pi = aY - bK$$

Taking account of lags, this becomes

$$\pi_t = aY_{t-1} - bK_{t-1}$$

where π_t refers to profits during period t, Y_{t-1} and K_{t-1} are income and capital stock of the previous period respectively and a and b are constants. Capital consumption allowances are expressed as

$$R_t = kK_{t-1}$$

The above equation shows that capital consumption allowances are a fraction (k) of capital stock (K_{t-1}).

Duesenberry's investment function is a modified version of the accelerator principle,

$$I_t = \alpha\, Y_{t-1} + \beta\, K_{t-1} \qquad \text{...(1)}$$

where investment in period t is a function of income (Y) and capital stock (K) of the previous period (t–1). The parameter (α) represents the effect of changes in income on investment, while the parameter (β) represents the influence of capital stock on investment working through both the

marginal efficiency of investment and profits.

Since the determinants of investment also affect consumption, the consumption function can be written as,

$$C_t = f\,(Y_{t-1} - \pi_{t-1} - R_{t-1} + d_t\,)$$

where d_t stands for dividend payments in period t. Since $\pi = f\,(Y, K)$, $R = kY$ and $d=f\,(\pi)$, these independent variables can be subsumed under Y and K. Thus

$$C_t = aY_{t-1} + bK_{t-1} \qquad \text{...(2)}$$

The parameter, a, in equation (2) is MPC and it also reflects increase in profits. This increase is reduced by the effect of profits on dividends and the effect of changes in dividends on consumption. The influence of changes in capital stock on consumption is reflected by the parameter b. This influence results from the influence of capital stock on profits through the influence of profits on dividends on consumption.

The capital stock is represented by the following equation which is an identity,

$$K_t = (1-k)\,K_{t-1} + I_t$$

It is derived as under :

$$K_t = K_{t-1} + (I_t - R_t\,)$$

$$R_t = kK_{t-1}$$

$$K_t = K_{t-1} + I_t - kK_{t-1}$$

$$K_t = (1-k)\,K_{t-1} + I_t \qquad [\because\ I_t = \alpha\, Y_{t-1} + \beta\, K_{t-1}]$$

The capital stock equation can be written as

$$K_t = (1-k)\,K_{t-1} + \alpha\, Y_{t-1} + \beta\, K_{t-1}$$

$$= [(1-k)\,K_{t-1} + \beta\, K_{t-1}] + \alpha\, Y_{t-1}$$

$$= K_{t-1}\,[(1-k) + \beta\,] + \alpha Y_{t-1}$$

Or $$K_{t-1} = K_{t-2}\,[(1-k) + \beta] + \alpha\, Y_{t-2} \qquad \text{.....(3)}$$

The national income identity can be written as

$$Y_t = I_t + C_t$$

$$= \alpha\, Y_{t-1} + \beta\, K_{t-1} + aY_{t-1} + bK_{t-1} \qquad [\text{From equations (1) and (2)}=$$

$$= [\,\alpha\, Y_{t-1} + aY_{t-1}\,] + [\,\beta\, K_{t-1} + bK_{t-1}\,]$$

$$= (\,\alpha + a)\, Y_{t-1} + (\beta + b)\, K_{t-1} \qquad \text{.....(4)}$$

By substituting the value of K_{t-1} in equation (4), we get

$$Y_t = (\alpha + a)\, Y_{t-1} + (\beta + b)\,\{\,K_{t-2}\,[(1-k)\,\beta\,] + \alpha\, Y_{t-2}\}$$

$$= (\alpha + a)\, Y_{t-1} + (\beta + b)\,\{a\, Y_{t-2} + [(1-k) + \beta\,]\, K_{t-2}\} \qquad \text{.....(5)}$$

Again, the national income equation (4) can be written as

$$Y_{t-1} = (\alpha + a)\, Y_{t-2} + (\beta + b)\, K_{t-2}$$

Or $$(\beta + b)\, K_{t-2} = Y_{t-1} - (\alpha + a)\, Y_{t-2}$$

Having obtained the value of K_{t-2}, substitute it in equation (5),

$$Y_t = (\alpha + a)\, Y_{t-1} + \alpha\,(\beta + b)\, Y_{t-2} + [(1-k) + \beta\,]\,\{Y_{t-1} - (\alpha + a)\, Y_{t-2}\}$$

By taking common factors, we have

$$Y_t = [(\alpha + a)\, Y_{t-1} + (1 - k + \beta)\, Y_{t-1}] + [\alpha\,(\beta + b)\, Y_{t-2} - (\alpha + a)\,(1-k + \beta)\, Y_{t-2}$$

$$= [(\alpha + a) + (1 - k + \beta)]\, Y_{t-1} + [\alpha\,(\beta + b) - (\alpha + a)\,(1-k + \beta)\, Y_{t-2} \qquad \text{...(6)}$$

This is a generalised version of a multiplier-accelerator process.

Next Duesenberry compares his formulation with a simple multiplier-accelerator process. In a simple model of the multiplier-accelerator, investment equation takes the following form

$$I_t = \alpha\, Y_{t-1} - K_{t-1}$$

whereas in his model, the investment equation is in this form

$$I_t = \alpha\, Y_{t-1} + \beta\, K_{t-1}$$

where the value of parameter $\beta = -1$ and investment is regarded as net investment so that $k = 0$. Depreciation is not considered. The influence of changes in capital stock on consumption is also neglected so that $b = 0$.

$$\alpha(\beta + b) = \alpha(-1)$$

$$Y_t = [(\alpha + a) + (1 - 0 - 1)]\, Y_{t-1} + [\alpha(-1 + 0) - (\alpha + a)(1 - 0 - 1)\, Y_{t-2}$$

Or $$\boxed{Y_t = (\alpha + a)\, Y_{t-1} + a(-1)\, Y_{t-2}} \quad \ldots(7)$$

The *a* (MPC) in equation (7) will be much smaller than MPC out of disposable income because it reflects the influence of changes in income on profits and business savings. Simultaneously, the *a* in the above equation will be much less than the average capital-output ratio which is the accelerator in simple multiplier-accelerator models. An increase, say, $100 in income, with capital stock constant, will increase the rate of business investment by an amount which is not much larger than the increase in business savings resulting from $100 increase in income. It will be only, say, $25. Thus an increase in income will have a smaller immediate effect on expenditure than would occur in a simple multiplier-accelerator model.

On the other hand, the negative effect of an increase in capital stock, with income constant, will be much smaller than in the simple multiplier-accelerator model. If there is an increase in business capital stock of say, $100, income being constant, it will reduce profits by a very small amount and will have correspondingly a small effect on business investment. But a part of the decline in business investment will be offset by a reduction in business saving. Such changes will reduce the effect on an increase in income on expenditure for some time because investment will decline slowly, as capital accumulates, provided there is no further increase in income. The system will therefore, be much more stable than a simple multiplier-accelerator system.

THE FINANCIAL THEORY OF INVESTMENT

The financial theory of investment has been developed by James Duesenberry. It is also known as the *cost of capital theory of investment*. The accelerator theories ignore the role of cost of capital in investment decision by the firm. They assume that the market rate of interest represents the cost of capial to the firm which does not change with the amount of investment it makes. It means that unlimited funds are available to the firm at the market rate of interest. In other words, the supply of

funds to the firm is very *elastic*. In reality, an unlimited supply of funds is not available to the firm in any time period at the market rate of interest. As more and more funds are required by it for investment spending, the cost of funds (rate of interest) rises. To finance investment spending, the firm may borrow in the market at whatever interest rate funds are available.

Sources of Funds

Actually, there are three sources of funds available to the firm for investment which are grouped under internal funds and external funds. These are : (1) Retained earnings which include undistributed profits after taxes and depreciation allowances are *internal funds.* (2) Borrowing from banks or through the bond market; and borrowing through equity financing or by issuing new stock (shares) in the stock market are the sources of *external funds*.

1. Retained Earnings. Retained earnings are the cheapest source of funds because the cost of using these funds is very low in the short run. There is no risk involved in spending these retained earnings or to repay debt. In fact, the cost of using these funds is the opportunity cost which is the return that the firm could obtain to repay debt or to buy the shares of other companies. The opportunity cost of internal funds will be less than the cost of external funds. When the firm lends these funds to other borrowers, it usually earns the market rate of interest. If it borrows funds from banks or through the bond market, it has to pay a higher interest rate. This difference in interest rate is the opportunity cost to the firm.

2. Borrowed Funds. When the firm needs funds more than the retained earnings, it borrows from the banks or through the bond market. The cost of borrowed funds (rate of interest) rises with the amount of borrowing. As the ratio of debt service to earnings from investment of funds rises, the marginal cost of borrowed funds rises. This is because the opportunity cost (risk) of not repaying debt increases.

3. Equity Issue. A third source is equity financing by issuing new shares in the stock market. The imputed cost of equity funds is more costly than the opportunity cost of retained earnings or borrowed funds. Duesenberry

points out that "the yield cost of equity finance is usually of the order of 7 to 10 percent for large firms. To this must be added floatation costs plus any reduction in the value of existing shares resulting from the issue. The differential is further increased by the differential tax treatment of bond and equity finance."*

Cost of Funds

The cost of capital to the firm will vary according to its source and how much funds it requires. Keeping these considerations in view, we construct the marginal cost of funds curve *MCF* in Figure 8 which shows the various sources of funds. The cost of funds is measured on the vertical axis and the amount of investment funds on the horizontal axis. Region *A* of the *MCF* curve shows financing done by the firm from retained profits (*RP*) and depreciation (*D*). In this region, the *MCF* curve is perfectly elastic which means the true cost of funds to the firm is equal to the market rate of interest. The opportunity cost of funds is the interest forgone which the firm could earn by investing its funds elsewhere. No risk factor is involved in this region. Region *B* represents funds borrowed by the firm from banks or through the bond market. The upward slope of the *MCF* curve shows that the market rate of interest for borrowed funds rises as their amount increases. But the sharp rise in the cost of borrowing is not only due to a rise in the market rate of interest but also due to the imputed risk of increased debt servicing by the firm. Region *C* represents equity financing. No imputed risk is involved in it because the firm is not required to pay dividends. The gradual upward slope of *MCF* curve is due

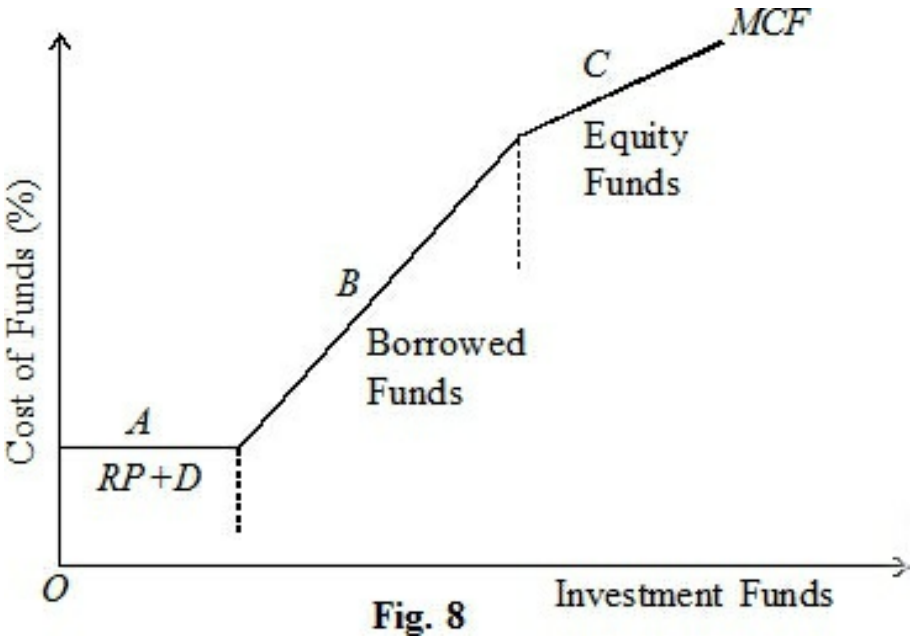


Fig. 8

to the fact that as the firm issues more and more of its stock, its market price will fall and the yield will rise.

The cost of funds may vary from firm to firm and consequently the shape and position of the *MCF* curve will differ from one firm to another. But in general, it will be like the *MCF* curve of Figure 8. If we aggregate *MCF* curves of different firms there will be a smooth S-shaped MCF_1 curve, as in Figure 9. This curve shifts upward from MCF_1 to MCF_2 when the cost of funds (interest rate) rises from R_1 to R_2 and shifts downward from MCF_2 to MCF_1 with the fall in the cost of funds from R_2 to R_1.

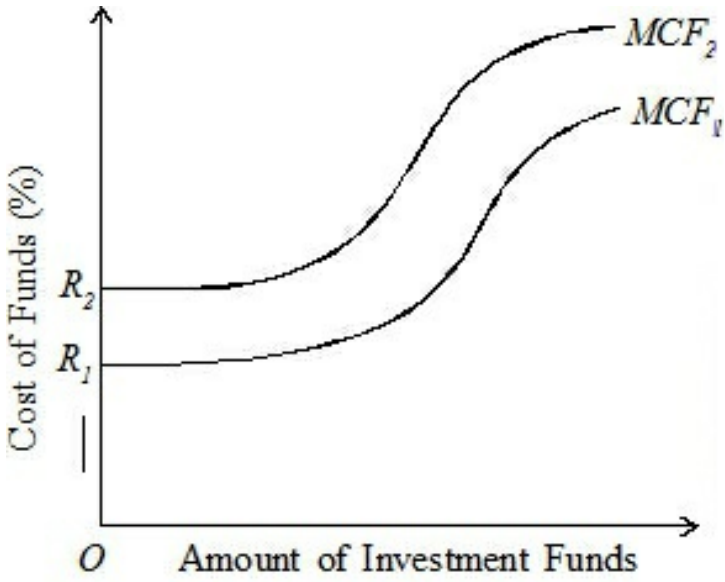


Fig. 9

---

* Interest on bonds and dividend from shares are taxed.

The amount of investment funds is determined by the intersection of *MEI* and *MCF* curves. The main determinants of the *MEI* curve are the rate of investment, output (income), level of capital stock and its age and rate of technical change. The determinants of *MCF* are retained earnings (profits minus dividends), depreciation, debt position of firms and market interest rate.

It is the shifts of the *MEI* and *MFC* curves that determine the level of investment funds. Suppose the *MEI* and *MCF* curves interest at point *E* in Figure 10 which determines *OI* investment at the interest rate (the cost of funds) *OR*. If the *MCF* curve shifts to the right to MCF_1 with the increase in retained earnings (profits) of the firm, the *MEI* curve will cut the MCF_1 curve at E_1. The cost of funds will fall from *OR* to OR_1, but investment funds will rise to OI_1, from *OI*. On the other hand, if the *MEI* curve shifts

to the right to MEI_1, with the increase in income and capital stock, it will cut the MCF_1 curve at point E_2. There will be increase in both the cost of funds to OR_2 and in the investment funds to OI_2.

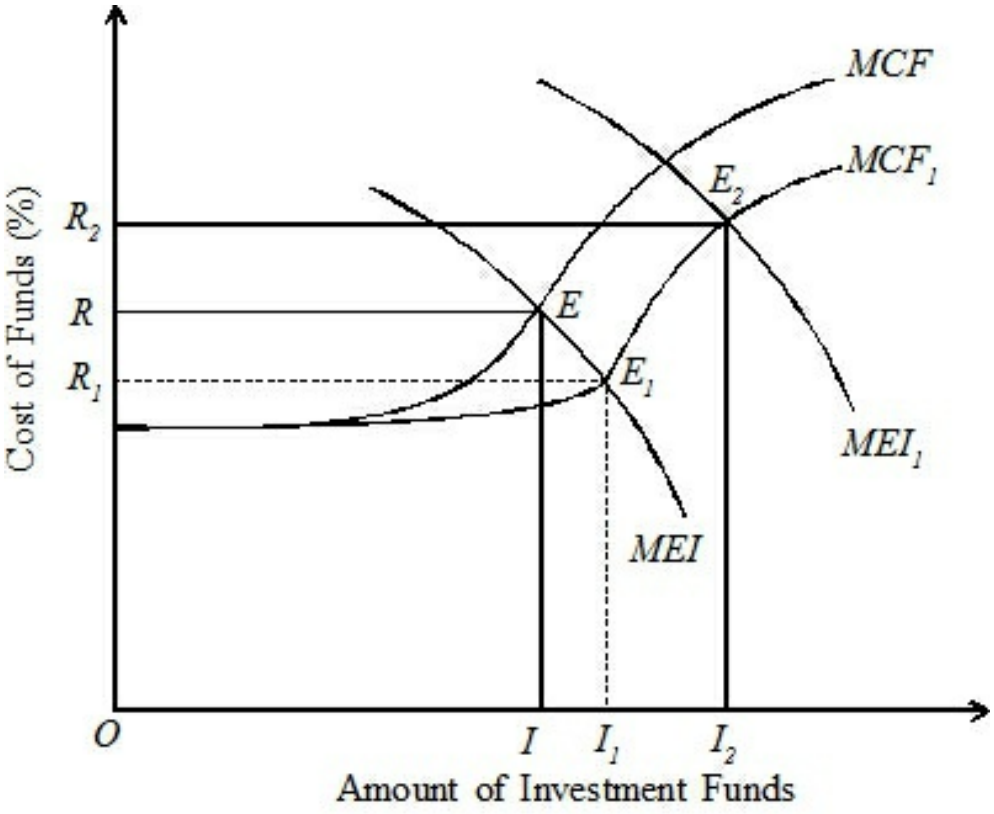


Fig. 10

The above explanation is related to the short-run behaviour of *MEI* and *MCF* curves. But the same factors that determine the position and shifts of these curves have different effects over the business cycle. Since the *MEI* curve depends primarily on output, it shifts backward to the left to MEI_1 when output (income) decreases in a recession, as shown in Figure 11. Both *MEI* and MEI_1 curves intersect the *MCF* curve in its perfectly elastic region. In a recession, retained profits decline but depreciation allowances remain with firms. So the elastic portion of the *MCF* curve becomes shorter. Meyer and Kuh[7] found that firms generally spend more of their retained earnings in recessions

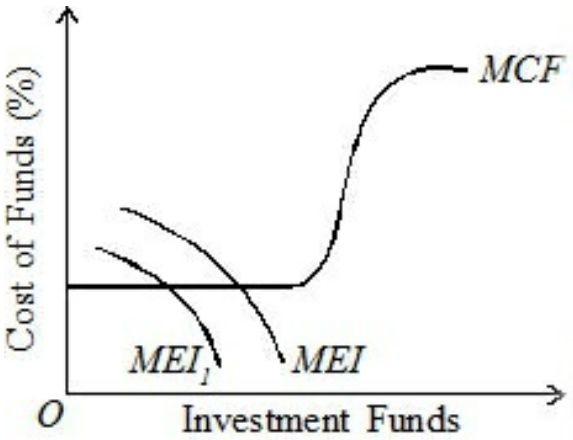


Fig. 11

and a low interest rate does not have any affect on investment. But when recovery starts, the MEI_1 curve shifts outward to the right to *MEI*. As a result, there is an increase in investment spending of the firm out of its retained earnings in the perfectly *elastic* portion of the *MCF* curve. Thus during a recession, monetary policy or the market rate of interest plays no role in determining the cost of capital of a firm.

---

7. *J.R. Meyer and E. Kuh, "Acceleration and Related Theories of Investment:An Empirical Inquiry".* R.E.&S, *Vol. 37, August, 1955.*

On the other hand, during a boom when output increases, the *MEI* curve shifts outward to the right to MEI_1, and intersects the *MCF* curve in its elastic rising region, as shown in Figure 12. In the upswing leading to boom, firms borrow funds on interest for investment spending. Thus monetary policy or interest rate is an important determinant of investment only in boom years.

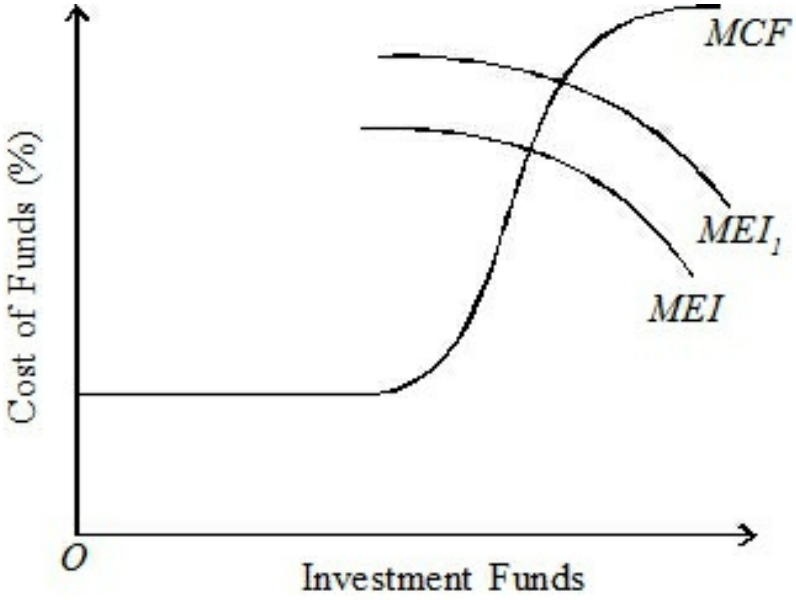


Fig. 12

Its Criticisms

The financial theory of investment has been criticised on the following grounds.

1. The results of studies by Meyer and Kuh on invesment behaviour of firms show that when demand is expanding rapidly, capacity expansion is the most important determinant of business investment during boom periods. In terms of our Figure 8, the *MEI* curve intersects the *MCF* curve in region *B*. In recessions and early years of recovery, the *MEI* curve shifts

back to region *A*, and the level of retained earnings provides the best explanation of investment spending.

2. Meyer and Kuh found that firms take a longer view while making investment spending, whereas Duesenberry explains a short-run model of investment. Their results indicate that firms primarily invest in capacity expansion during a boom period and their overall level of investment will not fall as much as indicated by Duesenberry's short-run model when the interest rate rises. On the other hand, firms generally spend most of their retained earnings on technological improvements to reduce costs and on advertisement to increase their market share.

3. Empirical evidence in the theory of investment by Kuh and Meyer[8] shows that monetary policy is the least effective of all the macroeconomic policy instruments. In the analysis represented in Figure 10, we have seen that the market rate of interest plays only a small role in the financial theory of investment. Critics point out that the main effect of rising interest rates would be to increase the steepness (or reduce the elasticity) of region *B* of the *MCF* curve. This would stop investment when retained earnings of firms had been exhausted. On the other hand, declining interest rates would flatten (increase the elasticity) region *B* of the *MCF* curve. This would have no effect in a recession if firms finance their investment spending from retained earnings. Thus monetary policy would be more effective in controlling a boom than in stimulating investment in recession.

4. This theory neglects the role of fiscal policy in investment which is more effective than monetary policy. A reduction in corporate taxes in a recession can increase investment by firms. On the other hand, an increase in corporate taxes can reduce investment and shift the *MCF* curve to the left. Changes in depreciation allowances can also help in manipulating investment in recessions and booms. Investment spending is also influenced by the level and changes in aggregate demand. Besides taxes, expenditure policy and other government measures also affect aggregate demand and the *MEI* curve which in turn influence the level of investment.

8. *E. Kuh and J.Meyer, "Investment Liquidity and Monetary Policy", in* Impacts of Monetary Policy, *1963.*

JORGENSONS' NEOCLASSICAL THEORY OF INVESTMENT

Jorgenson[9] has developed a neoclassical theory of investment. His theory of investment behaviour is based on the determination of the optimal capital stock. His investment equation has been derived from the profit maximisation theory of the firm.

Its Assumptions

Jorgenson's theory is based on the following assumptions:

1. The firm operates under perfect competition.

2. There is no uncertainty.

3. There are no adjustment costs.

4. There is full employment in the economy where prices of labour and capital are perfectly flexible.

5. There is a perfect financial market which means the firm can borrow or lend at a given rate of interest.

6. The production function relates output to the input of labour and capital.

7. Labour and capital are homogeneous inputs producing a homogeneous output.

8. Inputs are employed upto a point at which their *MPPs* are equal to their real unit costs.

9. There are diminishing returns to scale.

10. There is the existence of "putty-putty" capital which means that even after investment is made, it is instantly adapted without any costs to a

different technology.

11. The capital stock is fully utilised.

12. Changes in current prices always produce *ceteris paribus* proportional changes in future prices.

13. The price of capital goods equals the discounted value of the rental charges.

14. The firm maximises the present value of its current and future profits with perfect foresight in relation to all future values.

The Model

Jorgenson develops his theory of investment on the assumption that the firm maximises its present value. In order to explain the present value of the firm, he takes a production process with a single output (Q), a single variable input labour (L), and a single capital input (I-investment in durable goods), and p,w, and q representing their corresponding prices. The flow of net receipts (R) at time t is given by

$$R\,(t) = p\,(t)\,Q\,(\mathrm{t}) - w\,(t)\,L\,(t) - q(t)\,I(t) \qquad \ldots(1)$$

Where Q is output and p is its price; L is the flow of labour services and w the wage rate;I is investment and q is the price of capital goods.

The present value is defined as the integral of discounted net receipts which is represented as

---

9. *D.W. Jorgenson, "Capital Theory and Investment Behaviour,"* AER, *53, May 1963; and "The Theory of Investment Behaviour," in R.Ferber (ed.),* The Determinants of Investment Behaviour, *1967.*

$$W = \int_0^{\infty} e^{-rt} R(t)\,dt \qquad \ldots(2)$$

where W is the present value (net worth); e is the exponential used for continuous discounting; and r is the constant rate of interest.

The present value is maximised subject to two constraints. *First*, the rate of change of the flow of capital services is proportional to the flow of net investment. The constant of proportionality may be interpreted as the time rate of utilisation of capital stock, that is the number of units of capital service per unit of capital stock. Net investment is equal to total investment *less* replacement investment where replacement investment is proportional to capital stock. This constraint takes the form:

$$\dot{K}(t) = I(t) - \delta K(t) \qquad \text{...(3)}$$

where $\dot{K}(t)$ is the time rate of change of the flow of capital services at time (t) while δ is the rate of depreciation attached to capital stock. This constraint holds at each point of time so that K^*, K and I are functions of time. To sumplify the analysis, Duesenberry uses K in place of $K(t)$, I in place of $I(t)$, and so on.

Second, the levels of output and the levels of labour and capital services are constrained by a production function:

$$F(Q, L, K) = 0 \qquad \text{...(4)}$$

The marginal productivity of labour is equal to the real wage :

$$\frac{\partial Q}{\partial L} = \frac{w}{p} \qquad \text{...(5)}$$

Similarly, the marginal productivity of capital is equal to its real user cost:

$$\frac{\partial Q}{\partial K} = \frac{w}{p} \qquad \text{.....(6)}$$

where $c = q(r + \delta) - \dot{q}$ ...(7)

In the above equation, q is the average price of capital assets, r is the rate of discount, δ is the rate of depreciation of capital goods and q. is the rate of appreciation of capital assets or time derivative of q. Therefore, the crucial determinant of the optimal capital stock is c, the user cost of capital. Since most firms own rather than rent their their capital assets, therefore c is basically an implicit or shadow price constructed in order to

permit parallel analytical treatment of capital and labour inputs.

Equations (5) and (6) are called "myopic decision criteria" because the firm is engaged in a dynamic optimisation process and simply equates the *MP* of labour with the ratio of its price and *MP* of capital with the ratio of user cost of capital. There are two reasons for the myopic decision in the case of capital assets. *First*, it is due to the assumption of no adjustment costs so that the firm does not gain by delaying the acquisition of capital. *Second*, it is the result of the assumption that capital is homogeneous and it can be bought and sold or rented in a pertectly competitive market. The myopic decision is illustrated in Figure 13 where in the upper portion the two alternative time paths of output prices, P_1 and P_2, are shown and in the lower portion are shown the optimal capital stocks. In Panel (A), the output prices are identical up to time t_0, and then their time paths diverge when P_1 is always lower than P_2. With the myopic decision, the optimal capital stock are identical up to t_0 for both time path of output prices. But after that, for the time path of P_1 price, the optimcal capital stock K^*_1 moves at a constant rate, while for P_2 time path of output price, the optimal capital stock K^*_2 increases as the former rises. Thus in the Jorgenson model, there are no inter-temporal trade-offs.

Assuming that there are no adjustment costs, no uncertainty and perfect competition exists, as Jorgenson does, the firm will always be adjusted to the optimal capital stock so that $K=K^*$. Therefore, the question of adjustment to a discrete change in the interest rate does not rise. Instead, Jorgenson treats this problem as one of comparing two optimal paths of capital accumulation under two different interest rates. For this, he takes the demand for investment goods as given by the following equation :

$$I = \overset{*}{K} + \delta \qquad \text{...(8)}$$

where I stands for gross demand for investment goods, K^* the rate of change in capital stock, δ the rate of depreciation and K the fixed level of capital assests which is expressed as

$$K = f(w, c, p) \qquad \text{...(9)}$$

The condition of equation (9) implies that with *w* and *p* fixed, *c* must remain unchanged. From the expression for *c* in equation (7), this, in turn, implies that holding the price of investment goods constant, the rate of change of price of investment goods must vary as the interest rate varies so as to leave *c* unchanged.

Formally, this condition can be represented by

$$\frac{\partial c}{\partial r} = 0$$

where *r* is the interest rate.

This condition implies that the own-interest rate on investment goods $(r\text{-}q/\dot{q})$ must be left unchanged by variations in the interest rate.

Jorgenson assumes that all changes in the interest rate are exactly compensated by changes in the price of investment goods so as to leave the own-interest rate on investment goods unchanged. This condition implies that

$$\frac{\partial^2 q}{\partial t \partial r} = q$$

He further assumes that changes in the time path of interest rate leave the time path of forward or discounted prices of capital goods unchanged. This condition implies that

$$\frac{\partial^2 c}{\partial t \partial r} = c$$

Combining these two conditions, we obtain

$$\frac{\partial I}{\partial r} = \frac{\partial K}{\partial c} \times c < 0$$

It implies that the demand for investment goods in two alternative situations is a decreasing function of the interest rate. This is illustrated in

Figure 14 where in Panel (A), c_1 is the path of user cost of capital before a rise in the interest rate at t_0 time, and c_2 is the path after the change in interest rate. But c is constant at time t_0. Assuming other price p and w as given, K_1* is the path of optimal capital when the interest rate is unchanged, and K_2* is the path after the rise in interest rate. Thus at time t_0, a rise in the interest rate lowers the demand for investment goods. This is obtined by comparing two alternative and continuous paths of optimal capital accumulation.

Jorgenson concludes that the demand for investment goods depends on the interest rate by comparing two alternative and continuous paths of capital accumulation depending on a time path of the interest rate.

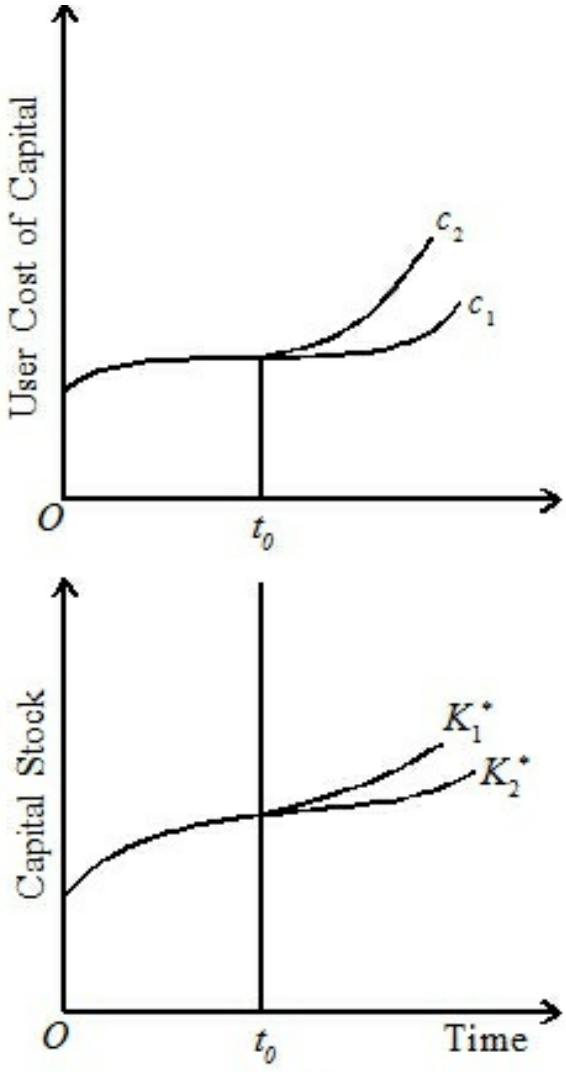


Fig. 14

Its Criticisms

Jorgenson's neoclassical theory of investment has been criticised on the following grounds :

1. Jorgenson derives his investment function from such assumptions which do not clarify how the actual capital stock adjusts to the optimal capital stock.

2. Jorgenson's theory is based on the assumption of full employment in the economy where prices of labour and capital are perfectly flexible so that producers and consumers can anticiapte changes in demand, supplies and prices of goods, But this is not a reality because there are long time lags for orders to be executed for capital goods which often lead to the fall in

investment demand and the consequent idle capacity and labour unemployment in both consumer and capital goods industries.

3. Jorgenson's analysis is based on expected quantities and prices that are perfectly foreseen. But foresight is never perfect. Moreover, Jorgenson does not proivde any mechanism for the formation of these expectations, except assuming that changes in current prices produce proportional changes in future prices. Further, he does not tell us anything about the expected future quantities to be sold.

4. The classical production function assumed by Jorgenson connects current investment with future outputs, and perfect foresight provides the exact current investment which produces the expected quantities of goods. Again, foresight is never perfect and current investment of capital may not be fully utilised in the future. Rather, there may be capital shortage in the future.

5. Jorgenson's definition of user cost is vague. It does not imply that future values of c (uses costs) will be identical. Consequently, a rise in the interest rate raises future user costs thereby lowering the future optimal path of capital accumulation than it otherwise would have been.

6. Jorgenson does not give a very clear economic account of his mathematical results.

7. Jorgenson labels his model as the neoclassical theory of investment but it seems to bear little relationship with the classical theory of investment.

TOBIN'S Q THEORY OF INVESTMENT

Nobel laureate economist James Tobin has proposed the *q* theory of investment which links a firm's investment decisions to fluctuations in the stock market. When a firm finances its capital for investment by issuing shares in the stock market, its share prices reflect the investment decisions of the firm. Firm's investment decisions depend on the following ratio, called *Tobin'sq*:

$$q = \frac{\text{Market Value of Capital Stock}}{\text{Replacement Cost of Capital}}$$

The market value of firm's capital stock in the numerator is the value of its capital as determined by the stock market. The replacement cost of firm's capital in the demominator is the actual cost of existing capital stock if it is purchased at today's price. Thus Tobin's *q* theory explains net investment by relating the market value of firm's financial assets (the market value of its shares) to the replacement cost of its real capital (shares).

According to Tobin, net investment would depend on whether *q* is greater than ($q>1$) or less than 1 ($q<1$). If $q>1$, the market value of the firms shares in the stock market is more than the replacement cost of its real capital, machinery etc. The firm can buy more capital and issue additional shares in the stock market. In this way, by selling new shares, the firm can earn profit and finance new investment. Conversely, if $q<1$, the market value of its shares is less than its replacement cost and the firm will not replace capital (machinery) as it wears out. Let us explain it with the help of an example. Suppose a firm raises finance for investment by issuing 10 lakh shares in the stock market at Rs 10 per share. Currently, their market value is Rs 20 per share. If the replacement cost of the firm's real capital is Rs 2 crores then the *q* ratio is 1.00 (= Rs 2 crores market value / Rs 2 crores replacement cost). Suppose the market value rises to Rs 40 per share . Now the *q* ratio is 2 (=Rs 40/Rs20) . Now the market value of its shares gives Rs 2 crores (=Rs 4 crores–Rs 2 crores) as profit to the firm. The firm raises its capital stock by issuing 5 lakh additional shares at Rs 40 per share. Rs 2 crores collected through the sale of 5 lakh shares are ulitised for financing new investment by the firm.

Panels (A) and (B) of Fig. 15 illustrate how an increase in Tobin's *q* induces a rise in the firm's new investment. It shows that an increase in the demand for shares raises their market value which raises the value of *q* and investment. The demand for capital is shown by the demand curve *D* in Panel (A). The relative value of *q* is taken as unity, as the market value and replacement cost of capital stock are assumed equal. The initial equilibrium is determined by the interaction of demand for capital and the

available supply of capital stock *OK* at point *E*, which is fixed in the short run.

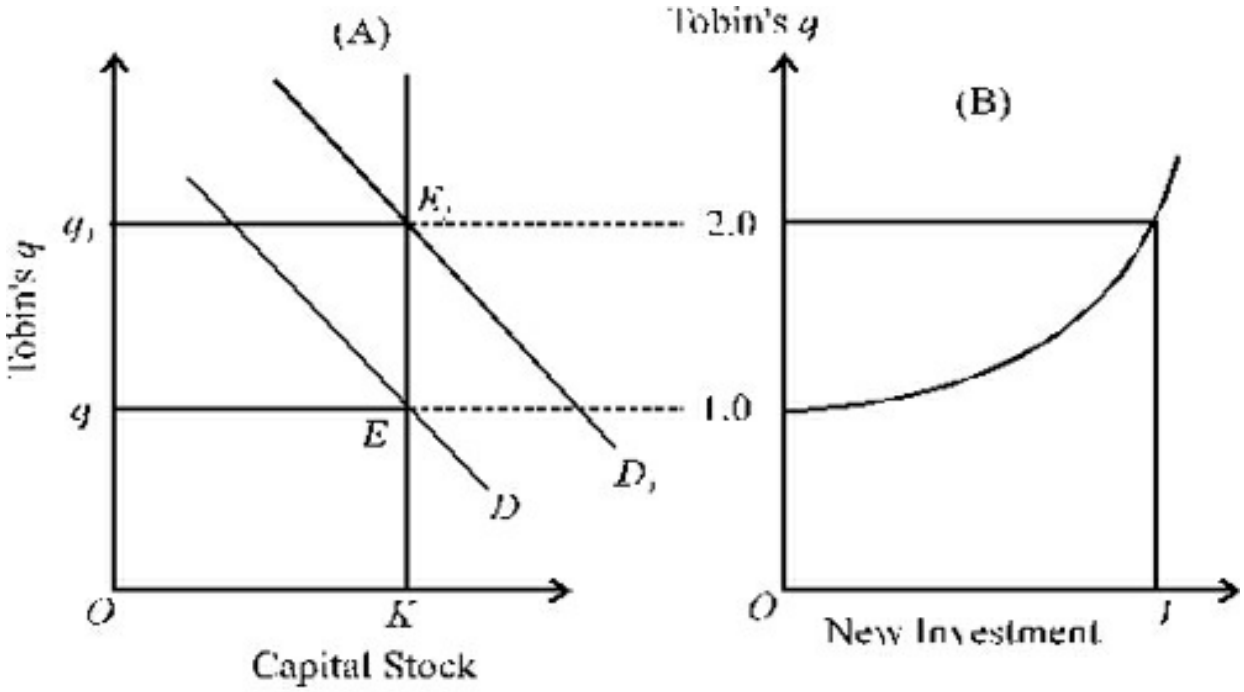


Fig. 15

The demand for capital depends mainly on two factors. *First,* the level of wealth of the people. The higher is the level of wealth, the more shares people wish to have in their wealth portfolio. *Second,* the real return on other assets such as government bonds or real estate. A fall in the real interest rate on government bonds would induce people to invest in shares than in other forms of wealth. This would increase the demand for capital and raise the market value of capital above its replacement cost. This means rise in the value of Tobin's *q* above unity. This is shown as the rightward shift of the demand curve to D_1. The new equilibrium is established at E_1 in the long run when the replacement cost rises and equals the market value of capital. The rise in the value of *q* to q_1 induces an increase in new investment to *OI,* as shown in Panel (B) of the figure.

Implications

Tobin's *q* theory of investment has important implications. Tobin's *q* ratio provides an incentive to invest for firms on the basis of the stock market. It not only reflects the current profitability of capital but also its expected

future profitability. Investment is expected to be higher in the future when the value of *q* is larger than 1. Tobin's *q* theory of investment induces firms to undertake net investment even when *q* is less than 1 in the present. They may adopt such economic policies which bring future profitability by raising the market value of their shares.

SELECTED READINGS

R. Dornbusch and S. Fisher, *Macroeconomics*, 1978.

P.N. Junankar, *Investment: Theories and Evidence*, 1972.

M.C.J. Surrey(ed.), *Macroeconomic Themes*, 1978.

Edward Shapiro, *Macroeconomic Analysis*, 1982.

EXERCISES

1. Summarise the accelerator theory of investment and provide examples of situations when this theory will not work.
2. Critically discuss the accelerator theory of investment.
3. Explain the flexible theory of investment.
4. Explain how adjustment is made when there are lags between levels of output and capital stock in the theory of accelerator.
5. Explain the profits theory of investment.
6. Explain the “cost of funds” theory of investment.
7. How has Duesenberry integrated the accelerator principle with the profits theory of investment?
8. Explain the financial theory of investment.
9. Critically examine the neoclassical theory of investment.
10. Explain Tobin's *q* of theory of investment.

CHAPTER 17

THE SAVING FUNCTION

MEANING OF SAVING FUNCTION

Saving is defined as the difference between disposable income and consumption : $S = Y - C$, where S is saving, Y is income and C is consumption. Thus the level of saving depends on the level of income. This is illustrated in Table 1.

Table 1: Relation Between Income and Saving

Income (Y)	*Consupm-tion (C)*	*Saving (S)*	*APS (S/Y)*	*MPS* $\Delta S/\Delta$
(1)	(2)	(3)	3/1 = (4)	(5)
0	20	–20	—	—
60	70	–10	—	—
120	120	0	0	—
180	170	10	0.6	0.17
240	220	20	0.8	0.17

Column (3) of the Table shows that when income is zero or very low, people dissave (minus Rs 20 crores or Rs 10 crores). They have to consume even if they are not earning or their consumption expenditure (Rs 70 crores) is more than their income (Rs 60 crores). When income (Rs 20 crores) equals consumption expenditure (Rs 120 crores), savings are zero. As income increases further by Rs 60 crores, their savings increase by Rs 10 crores. It shows that as income increases savings also increase

but by less than proportionately.

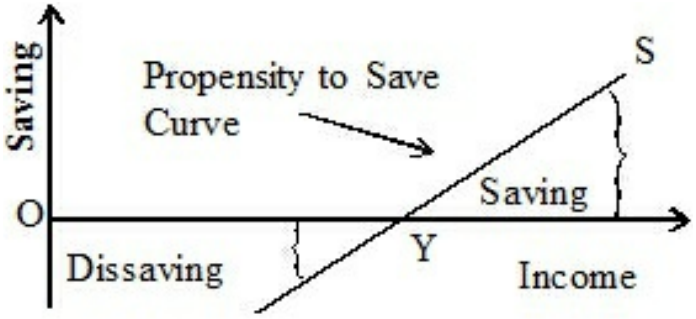


Fig. 1

This relation between saving and income is called the *propensity to save* or the *saving function.* It is represented as $S = f(Y)$. Thus the saving function indicates a functional relationship between *S* and *Y*, where *S* is the dependent and *Y* is the independent variable, that is, S is determined by *Y*. This relationship is based on the assumption "other things being equal" which means that all influences on savings are held constant and that income and saving increase by a constant amount, that is, income increases by Rs 60 crores and saving by Rs 10 crores, as shown in Table 1.

The propensity to save curve is shown in Fig. 1 where income is taken on the horizontal axis and saving on the vertical axis. The entire *S* curve with a definite position and slope is the propensity to save curve. The figure shows that below point *Y*, savings are negative because people dissave. At *Y*, savings are zero. Above *Y*, savings increase with the rise in income. The *S* curve is linear (straight line) because the rise in income and savings is at constant rates (Rs 60 crores and Rs 10 crores respectively).

The propensity to save is of two types : The average propensity to save and the marginal propensity to save, which we explain below.

The Average Propensity to Save (APS)

The *APS* is the ratio of saving to income. It is found by dividing saving by income, or $APS = S/Y$. It tells us about the proportion of each income level that people will save *i.e.*, they will not spend on consumption. For example, in Table 1 at an income level of Rs 180 crores, the consumption expenditure is Rs 170 crores and the savings are Rs 10 crores. The *APS* is 0.06 which means that people save 6 per cent of their income, as shown is column (4) of the Table. It is to be noted that as income increases, the average propensity to consume (*APC*) decreases from 0.94 to 0.92. But

APS increases from 0.06 to 0.08.*

Diagrammatically, the *APS* is any point on the *S* curve. In Fig. 2, point S_1, measures the *APS* of the *S* curve which is S_1Y_1/OY_1.

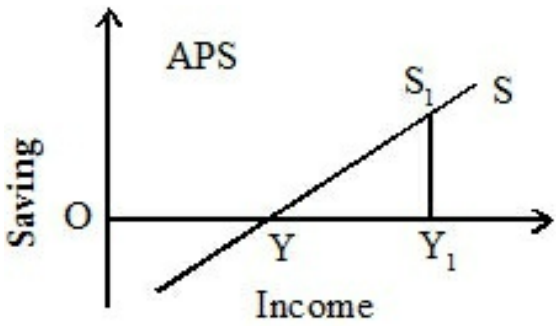


Fig. 2

* APC + APS must equal 1.

The Marginal Propensity to Save (MPS)

The *MPS* is the ratio of the change in saving to the change in income. It can also be defined as the rate of change in *APS* as income changes. It can be found by dividing a change in saving by a change in income, *i.e.*, $\Delta S/\Delta Y$. For example, in Table 1 when income increases from Rs 180 crores to Rs 240 crores, savings increase from Rs 10 crores to Rs. 20 crores so that ΔY = Rs. 60 (=240–180) crores and ΔS = Rs 10 (= 20 – 10) crores and the *MPS* = 10/60 = 0.17. It means that 17 per cent of income is saved, as shown in column (5) of the Table. It is constant at 0.17 because $\Delta S/\Delta Y$ = 10/60 is constant.

Diagrammatically, the *MPS* is measured by the gradient or slope of the *S* curve at a point or over a small range. This is shown in Fig. 3 by *AB*/*BC* where *AB* is the change in saving ΔS and *BC* is the change in income ΔY.

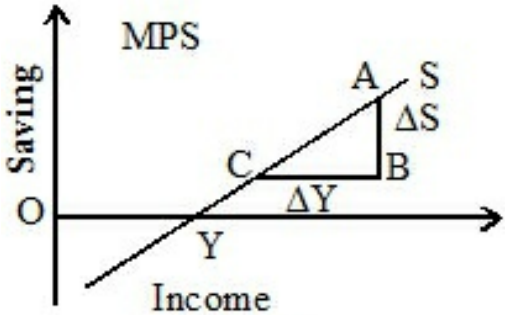


Fig. 3

Thus *APS* and *MPS* are two different concepts. The *APS* relates total saving to total income. On the other hand, the *MPS* relates a *change* in saving to a *change* in income.

Determinants of Savings

Savings depend upon the will to save, power to save and facilities to save. These three determinants of savings are discussed as under.

(a) Will to Save

A portion of the income can be saved only if a person has the will to save. Nothing can be saved by a person, if he does not want to save. This, in turn, depends upon the following factors.

1. Family Affection. It is the natural love and affection for the family for which people save. Every person has some affection for his children and family members. To make them enjoy life and to better their future prospects, he earns more and saves more. He wants to leave more property for them. For all this, he must have the will to save.

2. Precaution. Need for money can arise at any time in the future. So people keep money or wealth with them. People save in order to provide for old age, disease, accident and unforeseen needs and emergencies, etc. Therefore, every person wants to save as a precaution against future unforeseen needs.

3. Standard of Living. A person may want to raise his standard of living which can be achieved only through savings from his present income.

4. Farsightedness. Future is always uncertain. A farsighted person wants to make provision for education, marriage, etc. of his children. He wants to secure his old age. All this can be done by savings in the present time.

5. Calculating Mind. Certain persons have a calculating mind and they want to increase their future incomes. Therefore, they save out of their present incomes so as to earn more in future by investing the saved amount.

6. Enterprise. Persons doing trade or business want to take advantage of fluctuations in the rate of interest. They save more if they expect a rise in the rate of interest in the near future. Sometimes, savings are done to start a new business or to expand the existing one.

7. Independent. Every man wants to be self-sufficient or economically independent. He does not want to borrow from anybody in case he needs money in the future. He, therefore, saves out of his present income.

8. Social Status. Only wealthy persons are respected in the present society. Everybody wants to enjoy a higher social status. It strengthens the will to save.

9. Miserliness. There are people who save without any specific purpose. Miser persons save only to satisfy their desire for wealth.

(b) Power to Save

Power to save refers to the capacity to save. It means what remains after meeting the consumption expenditure out of the present income. In spite of his willingness to save, a man cannot save, if he does not have the power to save. After providing for his consumption expenditure, if more money remains with him, his power to save will be more. Thus the power to save depends upon both the levels of income and consumption. A person can save if his income exceeds consumption. If the gap between income and consumption is more, the power to save will increase. The power to save can be increased only by increasing the income, as there is less possibility of decreasing the consumption expenditure. Therefore, the power to save of the people in a country depends on their income or factors affecting their income. The following factors determine the power to save.

1. Size of National Income. Primarily, the power to save of the people depends upon the national income of the country. The higher the national income, the greater will be the power to save. The low national income is the main reason of the low power to save in India.

2. Natural Resources. The economic conditions and the income of a country depend upon its natural resources, *i.e.*, availability of land, water, minerals, etc. The greater use of these natural resources helps in increasing production. It increases income which further raises the power to save. But it should be noted that mere availability of natural resources does not raise the power to save, if these resources are not properly utilised. India is rich in natural resources, still its power to save is very low because we do not utilise these resources gainfully.

3. Trade. Both internal and foreign trade affect income and the power to

save. Income increases with the development of internal and foreign trade which, in turn, increase the power to save.

4. Industrial Development. Industrial development increases the power to save through the increase in income.

5. Agricultural Development. In countries like India where agriculture is the main occupation, power to save depends upon the development of agriculture also. Agricultural development increases income which results in an increase in the power to save.

6. Efficiency of Labour. An efficient labour in a country helps in increasing its production, resulting in an increase in income. It, in turn, gives rise to the power to save. A less efficient labour proves a hindrance in the way of power to save.

7. Distribution of Wealth and Income. The power to save depends upon the distribution of wealth and income in the country. An unequal distribution of wealth and income contributes to an increase in the power to save. Unequal distribution concentrates wealth in a few hands who can save more after meeting their consumption requirements.

(c) Facilities to Save

Savings depend not only on the will and the power to save, they also depend upon the facility to save. These facilities are :

1. Peace and Security. People can save only if their life and property are secure. They will not save if there is no peace and security of life and property.

2. Banking Facilities. An efficient and developed banking system facilitates savings. Savings are secure and beneficial in the form of cash, if these are deposited in the banks. Lack of banking facilities diminishes savings because in the absence of banking facilities money remains in the hands of people which is readily available for spending.

3. Taxation Policy. The taxation policy also influences savings in the

country. Progressive taxation reduces savings because tax rates increase with the rise in incomes. Similar is the case with wealth and succession taxes. People generally show a tendency to save less. They feel that most of their incomes will be taken away in the form of taxes and, therefore, they save less. On the contrary, expenditure tax encourages savings because to avoid this tax, people save more out of their income instead of spending it.

4. Value of Money. Facilities to save require stability in the value of money. The value of money decreases with an increase in prices. People save less for fear of fall in the money value. Stability in the price level or the value of money encourages savings.

5. Investment Opportunities. Investment opportunities encourage savings. Savings increase if there are sufficient opportunities to invest in trade and commerce. The development of stock and exchange markets also results in more savings. Scarce investment opportunities result in less savings.

6. Economic Policy of Government. Facilities to save are also influenced by the economic policy of the goverment. If the government wants to adopt the socialistic pattern of society, it would like to nationalise different industries and people will have a tendency to save less.

THE PARADOX OF THRIFT

The concept of "paradox of thrift" was first introduced by Bernard Mandeville in the *Fable of the Bees* in 1714. It was later recognised by several classical economists and became an integral part of the Keynesian economics.

Thriftiness is usually regarded as a virtue. An increase in thrift on the part of an individual leads to greater saving and wealth. It is also regarded a *public virtue* because if people consume less, then more resources can be devoted to producing capital goods which lead to increase in income, output and employment. According to Keynes, thrift is a public virtue only if the propensity to invest is equally high. Otherwise, thrift is a *public*

vice if the increase in the propensity to save is unaccompanied by increase in the propensity to invest, *i.e*., (autonomous investment).

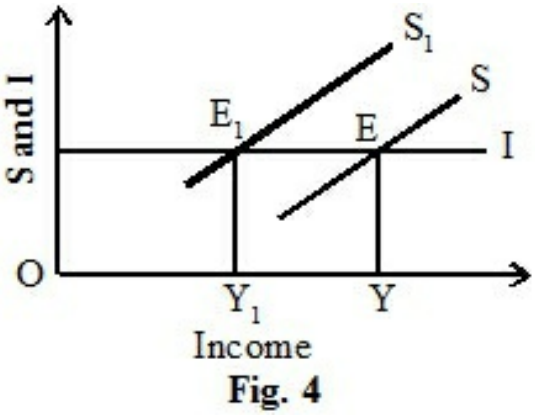


Fig. 4

Suppose people become thrifty and decide to save more out of a given level of income, given the propensity to invest. This will lead to a lower equilibrium level of income. This situation is illustrated in Fig. 4 where *S* is the saving curve and *I* is the income-inelastic investment curve. The equilibrium level of income is *E* where the two curves are equal. Suppose people become more thrifty. As a result, the saving curve shifts upward to S_1 with no change in the investment curve *I*. The new equilibrium point is E_1 which leads to a decline in the income level from *Y* to Y_1. Note that at the new equilibrium level Y_1, saving is the same as it was previously at *E*. (This is because both saving and investment have been taken along the vertical axis). Thus people's desire to save more has been frustrated. This is called the *paradox of thrift.*

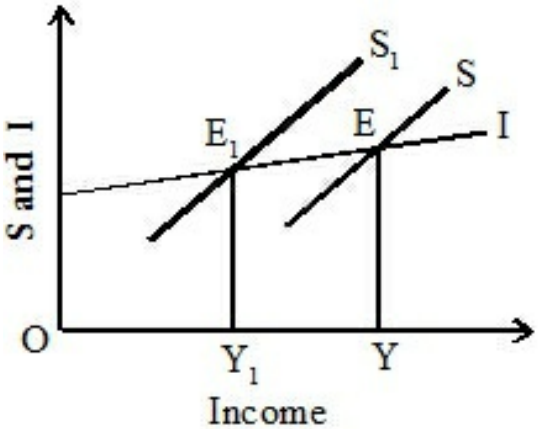


Fig. 5

The paradox of thrift can also be explained if investment is induced rather than autonomous. This is illustrated in Fig.5. The investment curve *I* slopes upward and intersects the *S* curve at point *E* where *Y* level of income is determined. Suppose people increase their savings due to thriftiness. As a result, the *S* curve shifts upward to S_1. It cuts the *I* curve at E_1 and the new equilibrium level of income is Y_1. The result is that there is reduction not only in the society's rate of saving but also in the rate of investment with the fall in income from *Y* to Y_1. People have tried to save more but they ended up saving less. This is the *paradox of thrift.*

Conclusion. The paradox of thrift highlights the fact that if people decide to save more, they end up saving less unless the increase in the propensity to save is offset by a higher propensity to invest, *i.e.*, the upward shift of the investment curve is larger than that of the saving curve. So saving is a virtue for a person or family because it leads to increase in saving and wealth. But it is a vice for the entire society because it leads to reductions in income, output and employment. Thus the paradox of thrift leads to the conclusion that saving is a private virtue and a public vice.

EXERCISES

1. What do you mean by saving ? Explain the determinants of savings.
2. Distinguish between the average propensity to save and the marginal propensity to save.
3. What do you mean by the saving function ? Explain the paradox of thrift.
4. Write a note on the paradox of thrift.
5. “Saving is a private virtue, but a public vice.” Discuss.

CHAPTER 18

SAVING AND INVESTMENT EQUALITY

All economists believe in the equality of saving and investment. But they differ as to the manner in which this equality is brought about. We study the following views on this problem.

THE CLASSICAL VIEW

The classicists believe in the existence of a fully employed economy where saving and investment are always equal. According to them, saving and investment are a function of the rate of interest. Algebraically, $S = f(R)$ and $I=f(R)$ where R is the rate of interest. Therefore, $S = I$. In a period of less than full employment, the inequality between the two is obvious which can be brought to equality through the mechanism of interest rate. When the rate of interest rises, saving rises and investment declines. On the other hand, with a fall in the rate of interest, saving falls and investment increases. If at any time, saving is less than investment, a rise in the rate of interest brings a decline in investment and increase in saving till saving equals investment. On the contrary, when saving is more than investment, the rate of interest falls, investment increases and saving declines till the two are equal at the new interest rate. This is explained in Figure 1, where

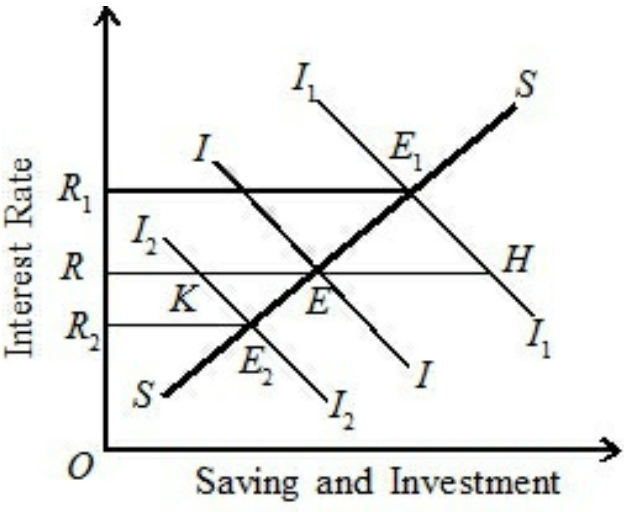


Fig. 1

saving and investment are measured on the horizontal axis and the rate of interest on the vertical axis. SS_1 is the saving curve which moves upward to the right with the rise in the interest rate, *II* is the investment curve. Both intersect at *E* when the rate of interest is *OR*. The economy is in full employment equilibrium because both saving and investment equal *RE*. Suppose investment increases (shown by I_1I_i curve) to *RH*, but saving is *RE* which is less than *RH* investment. The equality between the two can be brought about by raising the rate of interest to OR_1 where the SS_1 curve intersects the I_1 I_1 curve at E_1. On the contrary, if investment falls from *RE* to *RK* (shown by I_2 I_2 curve), saving *RE*>*RK* and a fall in the rate of interest to OR_2 brings the equality between saving and investment at E_2.

Keynes's Criticism of the Classical View

Keynes criticised the classical view about the saving-investment equality on the following grounds:

(*a*) Keynes does not agree with the classical view that the equality between saving and investment is brought about through the mechanism of the interest rate. According to him, it is changes in *income* which bring the two to equality, rather than the rate of interest.

(b) He also differs with the classicists that this equality via the rate of interest shifts only the investment curve and that the saving curve does not change. Keynes's view is that whenever the investment curve changes, there is a change in the saving curve also.

(c) Further, Keynes refutes the classical view that saving and investment are equal at the full employment level. He opines that since full employment is a rare phenomenon, saving-investment equality is found at less than full employment level.

THE KEYNESIAN VIEW

Keynes put forth two views with regard to the saving-investment equality. The first is the *accounting* or *definitional* equality between saving and

investment which is used in national income accounting. It tells us that *actual saving and actual investment are always equal at all times and at any level of income.* The second is the *functional* equality. In this sense, saving and investment are equal only at the *equilibrium level of income*. In other words, in the functional sense, saving and investment are not only *equal* but they are also in *equilibrium*. We study these views in detail below.

(1) The Accounting or Definitional Equality

Keynes wrote in his *General Theory* that saving and investment are necessarily equal in amount for the community as a whole, being different aspects of the same thing.[1] In order to show it, he defined saving and investment in such a manner as to establish their equality. Both saving and investment in the current period are defined as the excess of current income over current consumption $(Y_t - C_t)$, so that they are necessarily equal. Symbolically,

$$S_t = Y_t - C_t \qquad \ldots(1)$$

$$I_t = Y_t - C_t \qquad \ldots(2)$$

$\therefore$ $Y_t - C_t$ is common in equations (1) and (2), we have

$$S_t = I_t$$

---

1. J.M. Keynes, *op. cit*., p.74.

where S is saving, I is investment, Y is income, C is consumption and t is current period.

Keynes also established this equality in another way. He defined income in the current period (Y_t) as equal to current consumption (C_t) plus current investment (I_t); and saving in the current period (S_t) as the excess of current income over current consumption.

Thus $Y_t = C_t + I_t$ ...(i)

$Y_t = S_t + C_t$ ...(ii)

or $S_t = Y_t - C_t$

Therefore, from equations (*i*) and (*ii*), we have

$$C_t + I_t = S_t + C_t$$

or $$I_t = S_t$$

Thus saving and investment are the same thing. They are both the difference between income and consumption. So defined, they are always equal.

Criticism. This equality or rather identity between saving and investment which Keynes established in his *General Theory* has been severely criticised.

1. Relationship a Truism. According to Haberler, "If we accept these definitions *S* and *I* are necessarily equal over any period of time, because they are identically defined. It then becomes nonsensical to speak of, or to imply, differences between them." Ohlin has also described the relation between saving and investment as a case of equality *ex-definitions.* Thus this identity relationship is a *truism* and lacks an adjusting mechanism.

2. Decisions taken by Different People. It is difficult to understand how this equality can take place in actuality because saving and investment decisions are taken by different groups of people. Investment decisions are taken by businessmen on the basis of the rate of interest they have to pay for capital borrowed and the return they expect from it. On the other hand, decisions to save are based on different current and future requirements for saving by millions of individual families. Thus it is not possible for total saving to equal total investment in the economy except by coincidence.

3. Lagless Analysis. This accounting equality between saving and investment is a lagless analysis, which fails to explain the process of saving-investment relation. The saving-investment equality in this sense is an exercise in static analysis. Keynes failed to describe the actual dynamic

process of adjustment between saving and investment.

4 Does not distinguish between Financial Sources. according to Dr Lutz, the Keynesian definitions do not allow us to distinguish between investment financed out of inflationary credit or dishoarding from investment financed out of current supply of voluntary savings.

5. Not Ex-ante Equality. As pointed out by Ohlin, the equality between saving and investment is an ex-post equality and not an ex-ante one.

(2) The Functional Equality

The equality between saving and investment in the *functional* or *schedule* sense is brought about by the adjusting mechanism of income, as distinct from the classical view of variations in the rate of interest. In this sense, saving and investment are equal only at the equilibrium level of income. Income is functionally related to saving and investment. When saving is more than investment, income falls, and when investment is more than saving, income rises. This dynamic process of changes in income, saving and investment will continue till saving and investment are not only *equal* but are also in *equilibrium.* This is illustrated in the Table I given below.

Table I: Saving-Investment Equality

(Rs Crores)

Income	*Saving*	*Investment*	*Income Movements*
100	—15	10	
200	0	20	
300	15	30	*Expansion*
400	30	40	
500	45	50	
600	60	60	*Equilibrium*
700	75	70	
800	90	80	*Contraction*
900	105	90	

Table I shows that so long as investment is greater than saving income continues to increase till it reaches the income level of Rs 600 crores where saving and investment are in equilibrium at Rs 60 crores each. But after this point, saving exceeds investment and the equilibrium is reached when income contracts and again reaches equilibrium level of Rs 600 crores.

The equilibrium process of saving-investment equality is shown in Figure 2. Income is measured horizontally, and saving and investment vertically. SS is the saving curve and *II* the investment curve. When income is OY_1, investment is greater than saving, $I_1Y_1 > S_1Y_1$. Higher investment will lead to a rise in income and saving via the multiplier process till saving and investment are equal at the equilibrium income level *OY* represented by the point of interaction *E* of *II* and *SS* curves. When income is OY_2, saving is greater than investment, $S_2Y_2 > I_2Y_2$. This will bring a fall in income via the reverse operation of the multiplier till saving and investment are equilibrated at *OY* income level. Whenever there is a gap between saving and investment, income rises or falls till the equilibrium level is reached. Thus saving and investment are equal only in equilibrium.

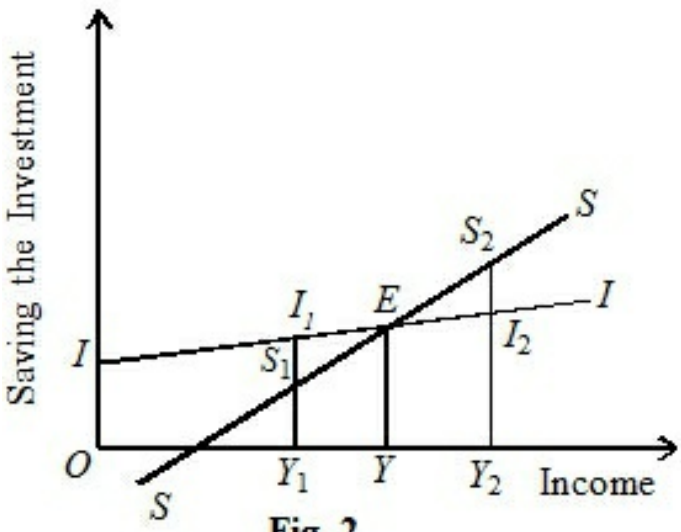


Fig. 2

Writing about the merits of this approach, Kurihara observes "that the Keynesian savings and investment schedules are to general equilibrium analysis what the Marshallian supply and demand curves are to partial equilibrium analysis. The modern income analysis makes use of the *functional* or schedule concepts of saving and investment in order to put life into dead figures, that is, to emphasize the *behaviour* of the economy as a whole instead of the *statistical result* of that behaviour."[2]

OTHER VIEWS

Some of the critics of Keynes especially Robertson and the Swedish economist Ohlin and others did not agree with him that saving and investment are necessarily equal.

The Robertsonian Approach

Robertson's approach involves period analysis. He defines saving as the difference between yesterday's income and current consumption expenditure, *i.e.* $S_t = Y_{t-1} - C_t$, where Y_{t-1} is yesterday's income. And current income consists of current consumption and current investment, *i.e.* $Y_t = C_t + I_t$. Thus Robertson defines investment in the Keynesian sense as expenditure on capital goods in the current period. It follows from the Robertsonian definitions of saving and investment that they are not necessarily equal. This can be shown symbolically as:

$$S_t = Y_{t-1} - C_t$$

$$Y_{t-1} = C_t + S_t$$

and

$$Y_t = C_t + I_t$$

or

$$Y_t - Y_{t-1} = C_t + I_t - C_t + S_t$$

$$= I_t - S_t$$

Thus current income exceeds yesterday's income provided current investment does not equal current saving. It is only when current investment equals current saving that current income would equal yesterday's income. Since current income cannot equal yesterday's income, therefore saving and investment are not necessarily equal.

Dr. Lutz regards Robertson's analysis more realistic than Keynes's because of the time lag involved between receipt and expenditure. The Robertsonian approach is dynamic as it is based on the period analysis whereas the Keynesian approach is static.

But Professor Klein does not agree with Lutz that the Robertsonian analysis of saving and investment is dynamic in the true sense. In his words: "Robertson's definitions are dynamic only in the most trivial sense.

They (definitions) define observed aggregate at different points of time, but they give no functional relation to economic behaviour so that one can solve for each variable in the system as a function of time alone. Robertson did not realise the existence of a consumption schedule, a saving schedule or an investment schedule."

The Swedish Approach

The Swedish economists Ohlin, Lundberg, Lindahl and Myrdal have defined saving and investment in *ex-ante* and *ex-post* senses. *Ex-ante* magnitudes are anticipated, desired, expected or planned, while *ex-post* magnitudes are actual or realised. Thus *ex-ante* saving and investment are those which people *plan or desire to make*. On the other hand, *ex-post* saving and investment are those which people *actually make*.

---

2. K. Kurihara, *op.cit*., pp. 76-77.

The Swedish economists point out that ex-post saving and investment are always equal. But ex-ante saving and investment are not necessarily equal because people who plan to save and to invest are different set of people. So long as savers and investors are influenced by different motives and belong to different groups, ex-ante saving and ex-ante investment are not necessarily equal. If ex-ante investment exceeds ex-ante saving, business activity will expand, and income will rise. Saving out of increased income will raise ex-post saving to the level of ex-post investment so that both are equal at higher level of income. Conversely, when ex-ante saving exceeds ex-ante investment, business activity will fall, and income will decline. Saving out of reduced income will reduce ex-post saving to the level of ex-post investment so that both are equal at a lower level of income. It is only when income is *constant* that ex-ante saving will equal ex-ante investment. And they are equal when the economy is in equilibrium. This is illustrated in Figure 3.

The equilibrium of the economy is shown at point *E* in Panel A of the figure where the aggregate demand curve *C+I* intersects the aggregate supply curve represented by the 45° line (*Y=C+I*), and *OY* level of income is determined. Corresponding to this income level, saving curve *S=f* (*Y*)

and investment curve *I* equal at point *R* in Panel B of the figure. Here ex-post saving equals ex-post investment. Suppose at this level, ex-ante investment exceeds ex-ante saving in the economy by *EF* in Panel A and by *RL* in Panel B of the figure. They are shown by the curves $C+I+\Delta I$ and $I+\Delta I$ respectively in the figure. As a result of increased ex-ante investment, business activity also rises. The level of employment, output, and income also increases to OY_1 when the aggregate supply curve 45° line intersects the new aggregate demand curve $C+I+\Delta I$ at point *G*. At this income level, ex-ante investment equals ex-ante saving and ex-post investment equals ex-post saving. This is shown in Panel B of the figure by point *M* where the rising saving curve $S=f(Y)$ intersects the increased investment curve $I+\Delta I$.

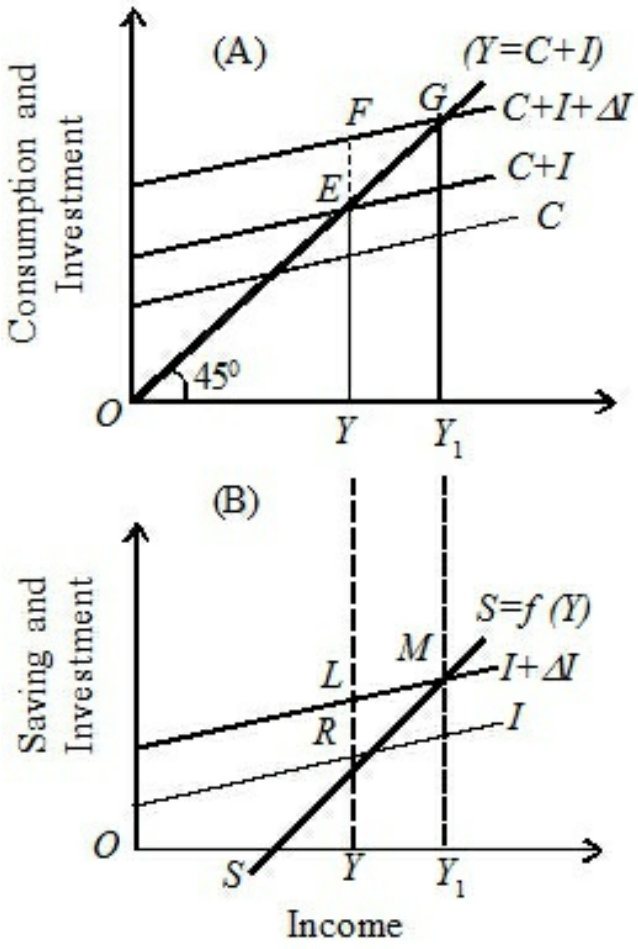


Fig. 3

We may conclude with Prof. Shapiro that *ex-post* saving and investment *can only be* equal and that *ex-ante* saving and *ex-ante* investment *can be* equal.

SELECTED READINGS

J.M. Keynes, *The General Theory of Employment, Interest and Money*, 1936

A.H. Hansen, *A Guide to Keynes*, 1953, Ch. *2*.

F.A. Lutz, "The Outcome of Saving-Investment Decisions", *Q.J.E.*, Aug.

1938.

L.R. Klein, *The Keynesian Revolution*, 1952.

EXERCISES

1. Critically examine the ex-ante and ex-post approaches to the inter-relations of saving and investment.
2. "Saving and investment are always equal." "Saving and investment are equal in equilibrium." Explain these statements.
3. Explain and reconcile the following statements:

   (*a*) Saving is always identical with investment; (*b*) Saving is equal to investment only when the economy is in equilibrium.

CHAPTER 19

THE MODEL OF NATIONAL INCOME DETERMINATION

INTRODUCTION

This chapter explains a general Keynesian model of national income determination in a capitalist economy. To simplify the analysis, it has been classified into a two-sector model, a three-sector model and a four-sector model. First two sectors are related to a closed economy in which there is no foreign trade and the last sector is concerned with the open economy.

TWO-SECTOR MODEL

A two-sector model of income determination of an economy consists only of domestic and business sectors.

Assumptions

The income determination in a closed economy is based on the following assumptions:

1. It is a two-sector economy where only consumption and investment expenditures take place. Thus the total output of the economy is the sum of consumption and investment expenditure.

2. Investment relates to *net* investment after deducting depreciation.

3. It is a closed economy in which there are no exports or imports.

4. There are no corporate firms in the economy so that there are no corporate undistributed profits.

5. There are no business taxes, no income taxes and no social security taxes so that disposable personal income equals NNP.

6. There are no transfer payments.

7. There is no government.

8. There is autonomous investment.

9. The economy is at less than full employment level of output.

10. The price level remains constant up to the level of full employment.

11. The money wage rate is constant.

12. There is stable consumption function.

13. The rate of interest is fixed.

14. The analysis relates to the short period.

Explanation

Given these assumptions, the equilibrium level of national income can be determined by the equality of aggregate demand and aggregate supply or by the equality of saving and investment.

Aggregate demand is the summation of consumption expenditure on newly produced consumer goods by households and on their services (C), and investment expenditure on newly produced capital goods and inventories by businessmen (*I*).

It is shown by the following identities:

$$Y = C+I \quad \text{....(1)}$$

Disposable Personal Income: $Y_d = C+S$ .....(2)

But $Y = Y_d$

$\therefore$ $C+I = C+S$

Or $I = S$

where Y = national income, Y_d = disposable income, C = consumption, S = saving, and I = investment.

In the above identities, $C+I$ relate to consumption and investment expenditures which represent aggregate demand of an economy. C is the consumption function which indicates the relation between income and consumption expenditure. The consumption function is shown by the slope of the C curve in Fig. l which is MPC (marginal propensity to consume). I is investment demand which is autonomous. When investment demand (I) is added to consumption function (C), the aggregate demand function becomes $C+I$.

$C+S$ identity is related to the aggregate supply of an economy. That is why, consumer goods and services are produced from total consumption expenditure and aggregate savings are invested in the production of capital goods.

In an economy, the equilibrium level of national income is determined by the equality of aggregate demand and aggregate supply ($C+I=C+S$) or by the equality of saving and investment ($S=I$).

We explain these two approaches one by one with the help of Figure 1 (A) and (B).

Equality of Aggregate Demand and Aggregate Supply

The equilibrium level of national income is determined at a point where the aggregate demand function (curve) intersects the aggregate supply function. The aggregate demand function is represented by $C+I$ in the figure. It is drawn by adding to the consumption function C the investment demand I. The 45° line represents the aggregate supply function, $Y = C+S$. The aggregate demand function $C+I$ intersects the aggregate supply

function $Y= C+S$ at point E in Panel *(A)* of Figure 1 and the equilibrium level of income OY is determined.

Suppose there is disequilibrium in aggregate supply and aggregate demand of the economy. Disequilibrium can be in either case, aggregate supply exceeding aggregate demand or aggregate demand exceeding aggregate supply. How will the equilibrium level of income be restored in the two situations ?

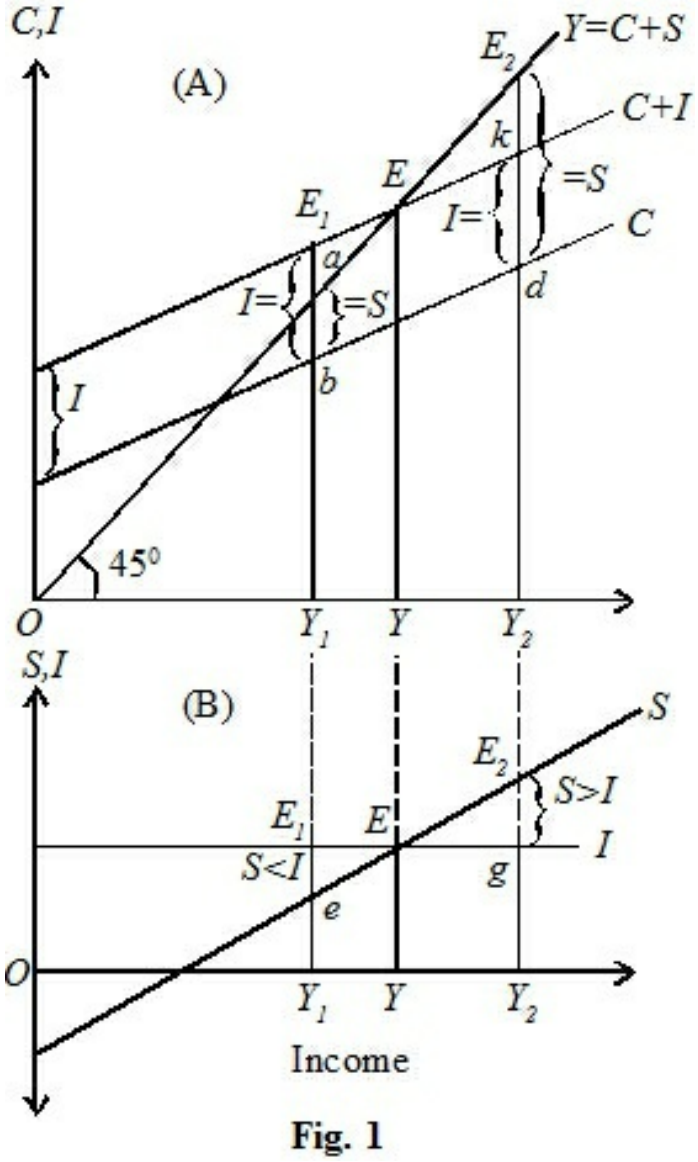


Fig. 1

First, take the case when aggregate supply exceeds aggregate demand. This is shown by OY_2 level of income in Panel (A) of the figure. Here aggregate output or supply is Y_2E_2 and aggregate demand is Y_2k. The disposable income is OY_2 $(=Y_2E_2)$. At this income level OY_2, consumers will spend Y_2d on consumption goods and save dE_2. But businessmen intend to make investment equal to dk in order to buy investment goods. Thus the aggregate demand for consumption goods and investment goods is $Y_2d + dk = Y_2k$. But aggregate supply (or output) Y_2E_2 is greater than aggregate demand Y_2k by kE_2 $(=Y_2E_2 - Y_2k)$. Therefore, the surplus output of goods worth kE_2 will be accumulated by businessmen in the form of unintended inventories. In order to avoid further inventory accumulation, they will reduce production. As a result of the reduction in output, income and employment will fall and the

equilibrium level of income will be restored at *OY* where the agrgregate supply equals aggregate demand at point *E*.

The second situation of disequilibrium when aggregate demand exceeds aggregate supply is shown by the income level of OY_1 in Panel (A) of the figure. Here the aggregate demand is Y_1E_1 and the aggregate output is Y_1a. The disposable income is OY_1 ($=Y_1a$). At this income level, consumers spend Y_1b on consumption goods and save *ba*. But businessmen intend to invest *bE*, to buy investment goods. Thus the aggregate demand is $Y_1b+bE_1=Y_1E_1$ which is greater than the aggregate supply of goods Y_1a by aE_1. To meet this excess demand worth aE_1, businessmen will have to reduce inventories by this amount. In order to stop further reduction in their inventories, businessmen will increase production. As a result of the increase in production, output, income and employment will increase in the economy and the equilibrium level of income *OY* will be restored again at point *E*.

Equality of Saving and Investment

The equilibrium level of income can also be shown by the equality of the saving and investment functions. Since the equilibrium level of income is determined when aggregate supply ($C+S$) equals aggregate demand ($C+I$) in the economy, intended (or planned) saving also equals intended (or planned) investment. This can be shown algebraically

$$C+S=C+I$$
$$\therefore \quad S=I$$

The equilibrium level of income in terms of the equality of saving and investment is shown in Panel (B) of Figure 1. where *I* is the autonomous investment function and *S* is the saving function. The saving and investment functions intersect at point *E* which determines the equilibrium level of income *OY*.

If there is disequilibrium in the sense of inequality between saving and investment, forces will operate in the economy and the equilibrium position will be restored. Suppose the income level is OY_2 which is above

the equilibrium income level *OY*. At this income level OY_2, saving exceeds investment by gE_2. It means that people are consuming and spending less. Thus aggregate demand is less than aggregate supply. This will lead to the accumulation of unintended inventories with businessmen. To avoid further accumulation of inventories, businessmen will reduce production. Consequently, output, income and employment will be reduced till the equilibrium level of income *OY* is reached at point *E* where *S=I*.

On the contrary, if the income level is less than the equilibrium level, investment exceeds saving. This is shown by OY_1 level of income when investment Y_1E_1 is greater than Y_1e saving. The excess of intended investment over intended saving means that aggregate demand is greater than aggregate supply by eE_1. Since aggregate output (or supply) is less than aggregate demand, businessmen will decrease inventories held by them. To stop further reduction in their inventories, they will increase production. Consequently, output, income and employment will increase in the economy and the equilibrium level of income *OY* will be again reached at point *E*.

The determination of equilibrium level of income simultaneously by the equality of aggregate demand and aggregate supply and of saving and investment is explained in Table I below.

Table 1

Panel (A)			*Panel (B)*
$Y=C+I$	at the equilibrium point	E	and $S=I$
$Y>C+I$	to the right of	E	and $S>I$
$Y<C+I$	to the left of	E	and $S<I$

THREE SECTOR MODEL

A three-sector model of income determination consists of a two-sector model and the government sector. The government increases aggregate demand by spending on goods and services, and by collecting taxes.

Government Expenditure

First, we take government expenditure. To explain it, given all the above assumptions except the government sector in the two-sector model, income determination is as follows :

By adding government expenditure (G) to equation (1) of the two-sector model, $Y=C+I$, we have

$$Y = C+I+G$$

Similarly, by adding government expenditure (G) to the saving and investment equation, when we have

$$Y = C+I+G$$
$$Y = C+S \qquad [\because S = Y-C]$$
$$\therefore \quad I+G = S$$

Both are illustrated in Figure 2(A) and (B). In Panel (A), $C+I+G$ is the new aggregate demand curve which intersects the aggregate supply curve 45° line at point E_1 where OY_1 is the equilibrium level of income. This income level is more than the income level OY without government expenditure. Similarly, according to the concept of saving and investment, the new investment curve $I+G$ intersects the saving curve S at point E_1 in Panel (B). Consequently, the income level OY_1 is determined which is more than the income level OY without government expenditure.

It should be noted that by adding government expenditure to consumption and investment expenditure ($C+I$), the national income increases by YY_1 which is more than the government expenditure, $\Delta Y > G$ in Panel (A) of the figure. This is due to the multiplier effect which depends upon the value of *MPC* or *MPS* where *MPC* or *MPS* < 1.

Taxation

Now we explain the effects of taxes on the level of national income. When the government imposes a tax, the amount of tax is reduced from the national income and what remains is the disposable income. Thus

$Y–T=Y_d$

where Y=national income, T=tax, and Y_d=disposable income. Now disposable income will be less than national income by the amount of tax, $Y_d<Y$. With the fall in disposable income, people will reduce expenditure on consumption. This will lead to reduction in national income, which will depend on the amount or rate of tax and the value of *MPC*.

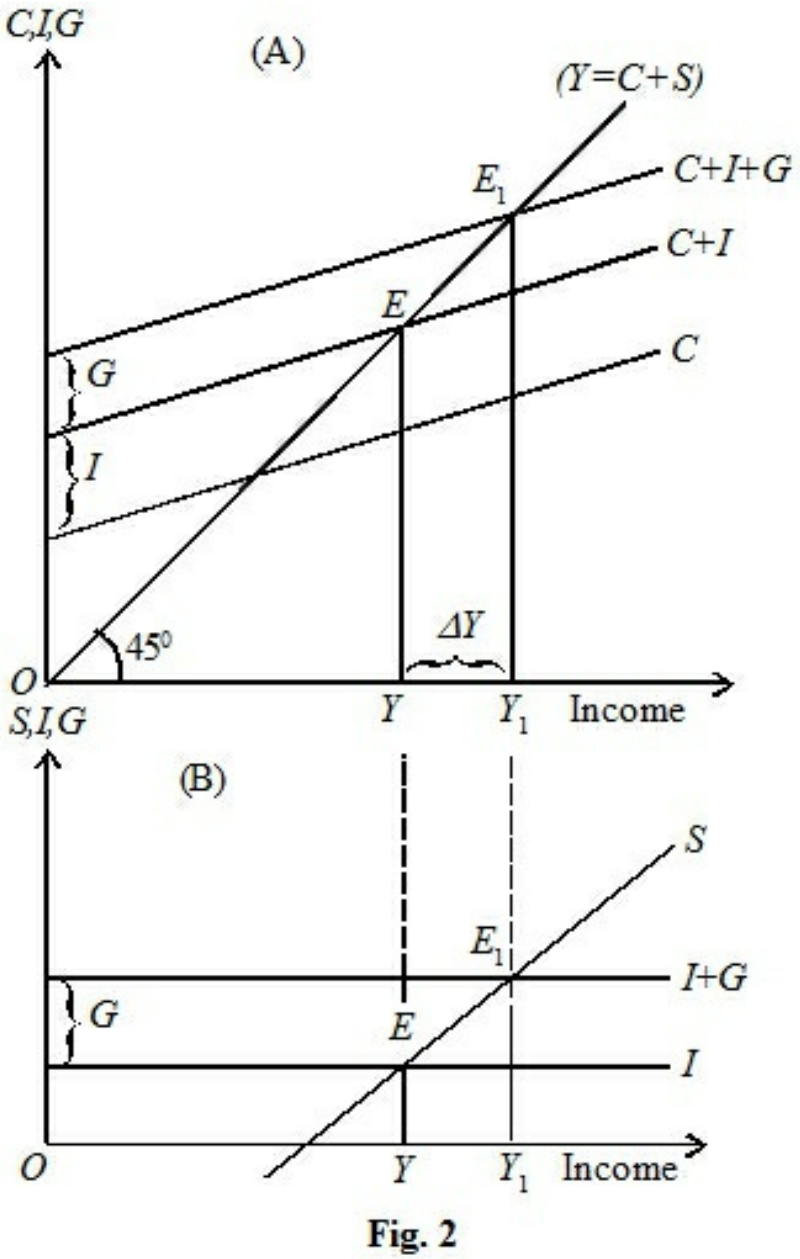


Fig. 2

Given all the above mentioned assumptions in which government expenditure is constant, the effects of taxes on national income are illustrated in the following figures.

First, the effect of a *lumpsum tax* on income is shown in Fig. 3. The equilibrium level of income without a tax is at point E where the aggregate demand curve ($C+I+G$) intersects the aggregate supply curve 45° line and the income level OY is determined. By imposing a lumpsum tax, the consumption function is reduced by the amount of tax. As a result, the aggregate demand curve $C+I+G$ shifts downwards to C_1+I+G and intersects the aggregate supply curve 45° line at point E_1. This results in

the reduction of income level from OY to OY_1. Thus with the imposition of a lumpsum tax, the national income is reduced by YY_1.

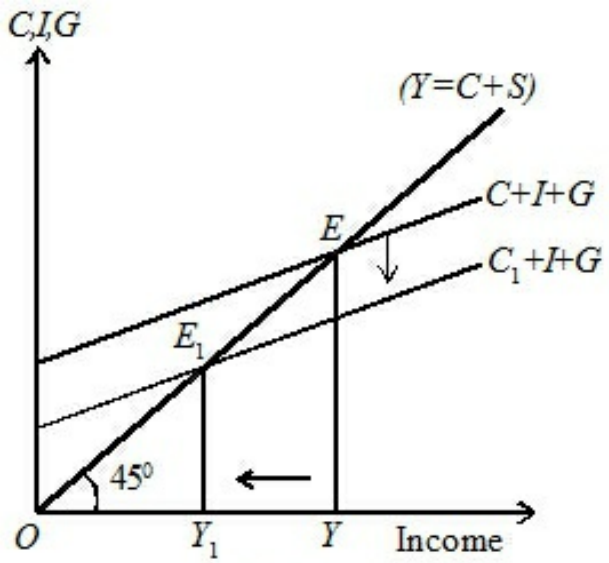


Fig. 3

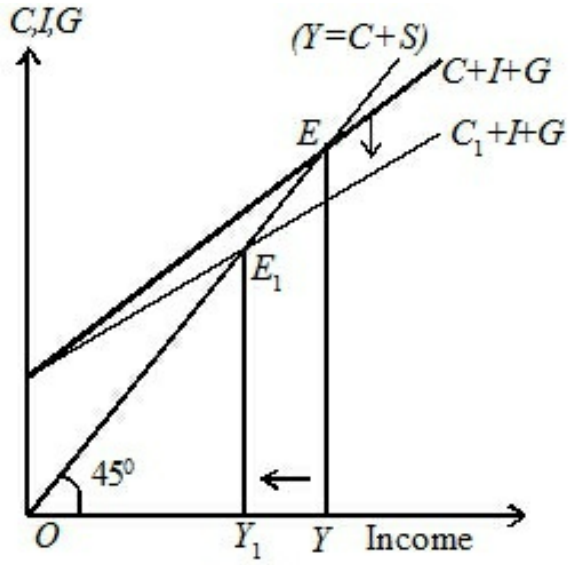


Fig. 4

Now we take a *proportional tax* which is imposed on income as a constant percentage. With the increase in the rate of tax, consumption and national income will decrease and vice versa. The effect of such a tax on income level is shown in Figure 4. The aggregate demand curve $C+I+G$ before the imposition of tax intersects the aggregate supply curve 45° line at point E and the income level OY is determined. After imposing the tax, the $C+I+G$ curve shifts downward to C_1+I+G due to a fall in consumption, and it intersects the 45° line at point E_1. Consequently, the equilibrium level of national income is reduced by YY_1.

Effect on Saving and Investment

The effect of a tax on saving and investment also determines the equilibrium of national income as follows:

$$Y=C+I+G$$

And $$Y=C+S+T$$

$\therefore$ $$Y=C+I+G=C+S+T$$

Or $$Y=I+G=S+T$$

It is clear from the above equation that when planned investment (*I*) *plus* government expenditure on goods and services (*G*) equal planned saving (*S*) *plus* tax (*T*), the equilibrium of national income is established. *I*+*G* are inflows or injections in the national income and *S*+*T* are outflows or leakages. If they are equal to each other, the national income is in equilibrium. This is shown in Fig. 5. Here, *E* is the equilibrium point before imposing the tax where *S* and *I*+*G* curves intersect and the income level *OY* is determined. With the imposition of a tax, the *S* curve shifts upward to the left as *S*+*T* and the new equilibrium is established at point E_1 with *I*+*G* and the national income falls from *OY* to OY_1.

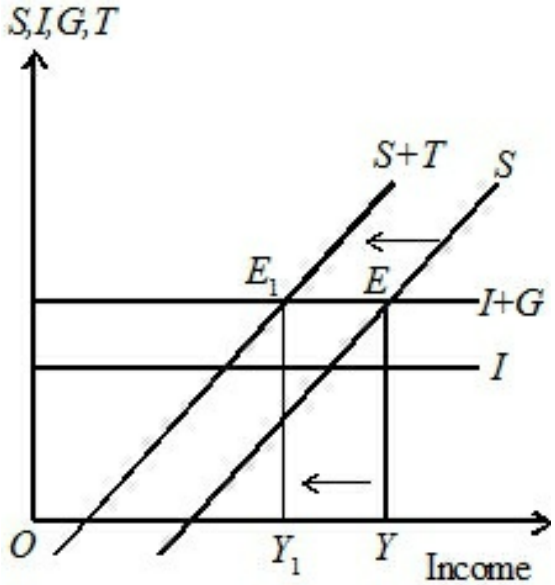


Fig. 5

FOUR SECTOR MODEL : INCOME DETERMINATION IN OPEN ECONOMY

We shall now show how national income is determined in an open economy. For this, we relax the assumptions that there are no exports or imports and government expenditures. This means that we shall have to add imports and exports and government expenditures and taxation in our analysis. It may be noted that government expenditures are like investment because they raise the demand for goods. They are *injections* in the national income. On the other hand, taxes are *leakages* in the national income like savings because they tend to reduce the demand for consumer goods. The impact of exports and imports is similar to that of the government expenditure. Exports are *injections* because they increase the demand for goods in the *same economy*. Imports, on the other hand, are *leakages* in the national income because they represent the supply of goods to the given economy.

Assumptions

The analysis of the determination of income in an open economy is based

on the following assumptions:

1. The domestic economy's international trade is small relative to total world trade.

2. There is less than full employment in the economy.

3. The general price level is constant upto the full employment level.

4. Exchange rates are fixed.

5. There are no tariffs, trade and exchange restrictions.

6. Gross exports are determined by external factors.

7. Exports (*X*), investment (*I*) and goverment expenditure (*G*) are autonomous.

8. Consumption (*C*), imports (*M*), savings (*S*) and taxes (*I*) are each a fixed proportion of national income (*Y*) and their relationships with national income are linear.

Determination of Equilibrium Level of Income

Given these assumptions, an open economy is in equilibrium when its national expenditure (*E*) is equal to its national income (*Y*). This can be shown in the following equation for the equilibrium level of income:

$$Y=E=C+I+G+(X-M)$$

But $Y=C+S+T$

$\therefore \quad C+S+T=C+I+G+(X-M)$

In the above analysis, $C+S+T$ is gross national income (GNI) and $C+I+G+$ (X-M) is gross national expenditure (*GNE*). Thus the equilibrium level of income in an economy is determined when aggregate supply, $GNI=GNE$, aggregate demand, or, $C+S+T=C+I+G+(X-M)$. This is shown in Figure 6 where C is the consumption function. On this curve, *I* autonomous investment is superimposed to form the $C+I$ function, and autonomous

government expenditure G is superimposed on $C+I$ to form the $C+I+G$ function. When net exports of $X–M$ are superimposed on $C+I+G$, we get the aggregate demand function $C+I+G+(X–M)$. The 45° line is the aggregate supply function which represents $C+S+T$.

It should be noted that so long as $C+I+G+$

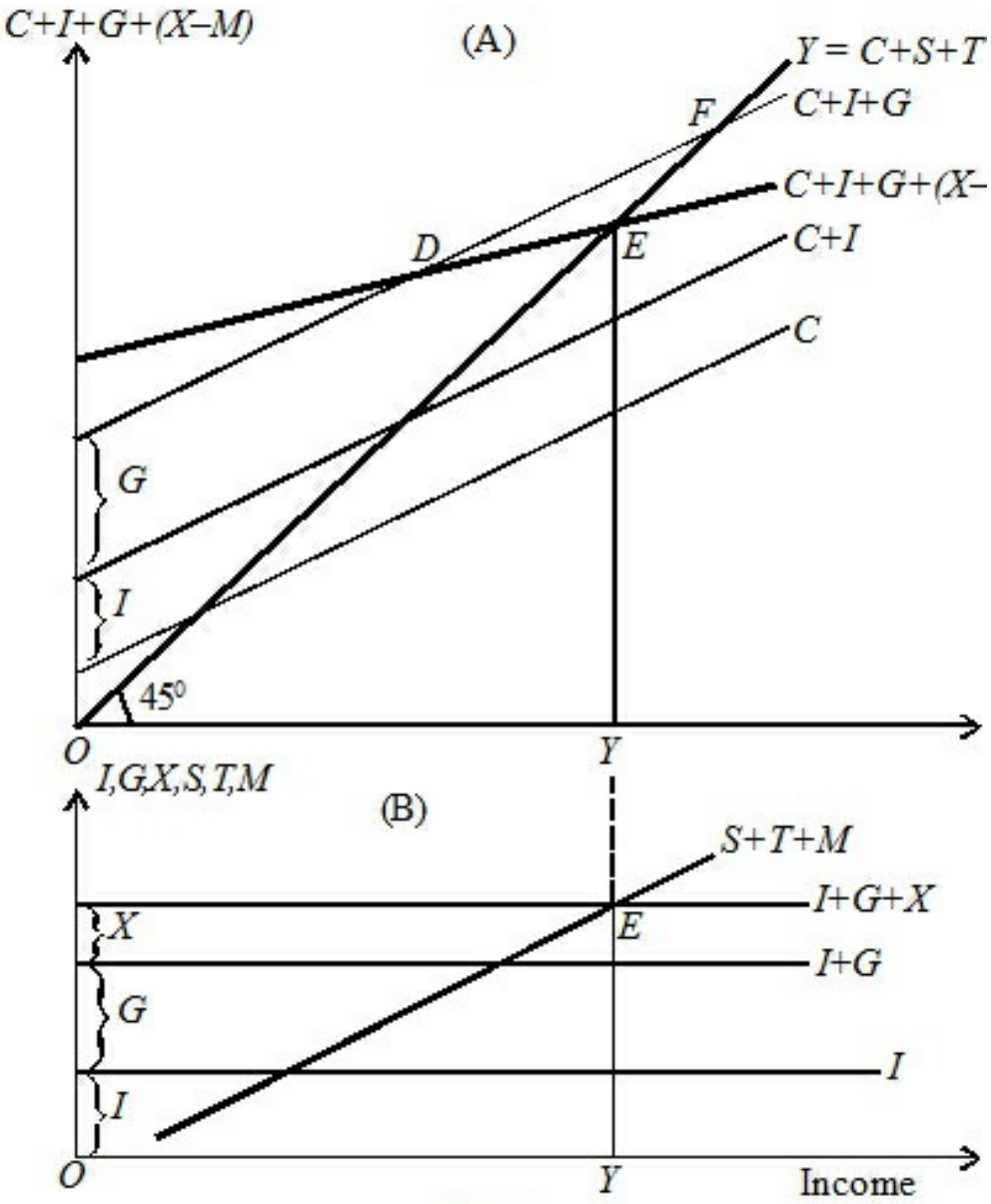


Fig. 6

$(X–M)>C+I+G$, exports exceed imports and there is net addition to aggregate demand. At point D in Panel (A) of the figure, $X–M=O$. Beyond point D,$C+I+G>C+I+G+(X–M)$ and imports exceed exports, and this gap continues to grow as income increases. This leads to net reduction in aggregate demand so that the aggregate demand function $C+I+G+(X–M)$

lies below the domestic demand function $C+I+G$.

The equilibrium level of income in an open economy, OY is determined at point E where the aggregate demand function $C+I+G+(X-M)$ intersects the aggregate supply function $C+S+T$.

This analysis shows that in the absence of foreign trade, the equilibrium level of income would have been at a higher level, as determined by the equality of $C+I+G=C+S+T$ at point F whereas with foreign trade it is at a lower point E.

There is also an alternative method for determining the equilibrium level of income in an open economy in terms of saving and investment equality.

Accordingly,

$$C+S+T=C+I+G+(X-M)$$

Or $$S+T=I+G+(X-M)$$

Or $$S+T+M=I+G+X$$

where $S+T+M$ refers to total income and $I+G+X$ to total expenditure. When $S+T+M$ is equal to $I+G+X$, the equilibrium level of income is determined. This is shown in Panel (B) of Fig. 6 where the $S+T+M$ curve intersects the $I+G+X$ curve at point E and the equilibrium level of income OY is determined.

EXERCISES

1. What is meant by equilibrium level of national income ? Show that this level is attained only when saving is equal to investment.
2. Explain the theory of income determination in a three sector model.
3. Explain income determination in a closed economy.
4. How is income level determined in an open economy ?
5. Explain income determination with government expenditure and taxation.

CHAPTER 20

THE KEYNESIAN THEORY OF INCOME, OUTPUT AND EMPLOYMENT: A SUMMARY

INTRODUCTION

Keynes restated his theory of employment in Chapter 18 of *General Theory*. We have analysed in detail the various dependent and independent variables of the Keynesian system in the foregoing chapters. Now we combine them together, as Keynes did, and present the theory of income, output and employment briefly.

THE KEYNESIAN THEORY OF INCOME, OUTPUT AND EMPLOYMENT

In the Keynesian theory, employment depends upon effective demand. Effective demand results in output. Output creates income. Income provides employment. Since Keynes assumes all these four quantities, viz., effective demand (*ED*), output (*Q*), income (*Y*) and employment (*N*) equal to each other, he regards employment as a function of income.

Effective demand is determined by two factors, the aggregate supply function and the aggregate demand function. The aggregate supply function depends on physical or technical conditions of production which do not change in the short-run. Since Keynes assumes the aggregate supply function to be stable, he concentrates his entire attention upon the aggregate demand function to fight depression and unemployment. Thus employment depends on aggregae demand which in turn is determined by consumption demand and investment demand.

According to Keynes, employment can be increased by increasing consumption and/or investment. Consumption depends on income $C(Y)$ and when income rises, consumption also rises but not as much as income. In other words, as income rises, saving rises. Consumption can be increased by raising the propensity to consume in order to increase income and employment. But the propensity to consume depends upon the psychology of the people, their tastes, habits, wants and the social structure which determine the distribution of income. All these elements remain constant during the short-run. Therefore, the propensity to consume is stable. Employment thus depends on investment and it varies in the same direction as the volume of investment.

Investment, in turn, depends on the rate of interest and the marginal efficiency of capital (*MEC*). Investment can be increased by a fall in the rate of interest and/or a rise in the *MEC*. The *MEC* depends on the supply price of capital assets and their prospective yield. It can be raised when the supply price of capital assets falls or their prospective yield increases. Since the supply price of capital assets is stable in the short-run, it is difficult to lower it. The second determinant of *MEC* is the prospective yield of capital assets which depends on the expectations of yields on the part of businessmen. It is again a psychological factor which cannot be depended upon to increase the *MEC* to raise investment. Thus there is little scope for increasing investment by raising the *MEC*.

The other determinant of investment is the rate of interest. Investment and employment can be increased by lowering the rate of interest. The rate of interest is determined by the demand for money and the supply of money. On the demand side is the liquidity preference (*LP*) schedule. The higher the liquidity preference, the higher is the rate of interest that will have to be paid to cash holders to induce them to part with their liquid assets, and vice versa. People hold money (*M*) in cash for three motives: transactions, precautionary and speculative. The transactions and precautionary motives (*M*) are *income elastic*. Thus the amount held under these two motives (M_1) is a function (L_1) of the level of income (Y), *i.e.* $M_1=L_1(Y)$. But the money held for speculative motive (M_2) is a function of the rate of interest (r), *i.e.* $M_2=L_2\ (r)$. The higher the rate of interest, the lower the demand for money, and vice versa. Since *LP* depends on the psychological attitude

to liquidity on the part of speculators with regard to future interest rates, it is not possible to lower the liquidity preference in order to bring down the rate of interest. The other determinant of interest rate is the supply of money which is assumed to be fixed by the monetary authority during the short-run.

The relation between interest rate, *MEC* and investment is shown in Figure 1, where in Panels (*A*) and (*B*) the total demand for money is measured along the horizontal axis from *M* onward. The transactions (and precautionary) demand is given by the L_1 curve at OY_1 and OY_2 levels of income in Panel (A) of the figure. Thus at OY_1 income level, the transactions demand is given by OM_1 and at OY_2 level of income it is OM_2. In Panel (B), the L_2 curve represents the speculative demand for money as a function of the rate of interest. When the rate of interest is R_2, the speculative demand for money is MM_2. With the fall in the rate of interest to R_1, the speculative demand for money increases to MM_1. Panel (C) shows investment as a function of the rate of interest and the *MEC*. Given the *MEC*, when the rate of interest is R_2, the level of investment is OI_1. But when the rate of interest falls to R_1, investment increases to OI_2.

In the Keynesian analysis, the equilibrium level of employment and income is determined at the point of equality between saving and investment. Saving is a function of income, *i.e.* $S=f(Y)$. It is defined as the excess of income over consumption, $S=Y-C$, and income is equal to consumption plus investment.

$$\begin{aligned} \text{Thus} \quad & Y = C+I \\ \text{or} \quad & Y-C = I \\ \because \quad & Y-C = S \\ \therefore \quad & I = S \end{aligned}$$

So the equilibrium level of income is established where saving equals investment. This is shown in Panel (D) of Figure 1 where the horizontal axis from *O* toward the right represents investment and saving, and *OY* axis represents income. *S* is the saving curve. The line I_1E_1 is the investment curve (imagine that it can be extended beyond *E* as in an *S* and

I diagram) which touches the *S* curve at E_1* Thus OY_1 is the equilibrium level of employment and income. This is the level of underemployment equilibrium, according to Keynes. If OY_2 is assumed to be the full employment level of income then the equality between saving and investment will take place at E_2 where I_2 E_2 investment equals Y_2E_2 saving.

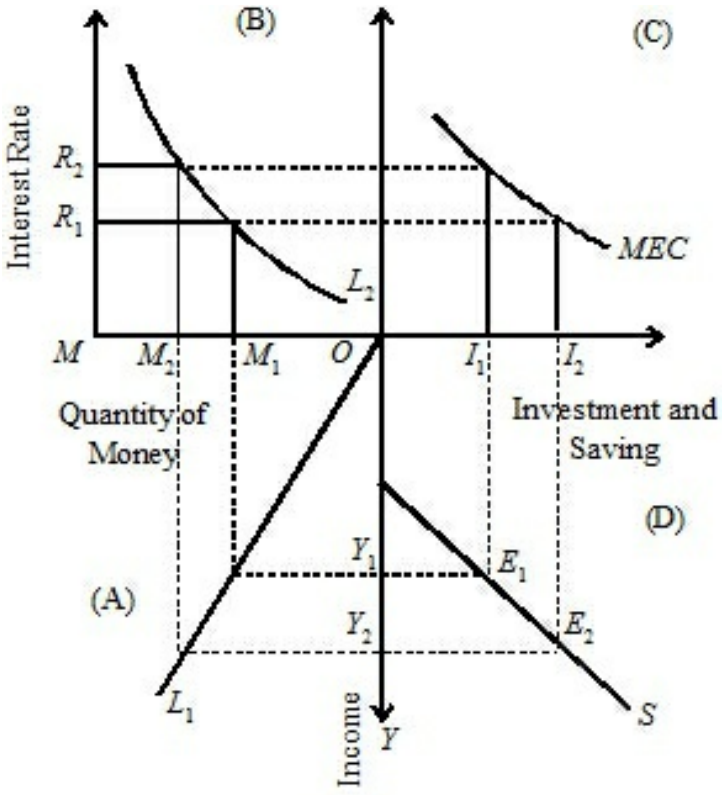


Fig. 1

The Keynesian theory of employment and income is also explained in terms of the equality of aggregate supply (*C*+*S*) and aggregate demand (*C*+*I*). Since unemployment results from the deficiency of aggregate demand, employment and income can be increased by increasing aggregate demand. Assuming the propensity to consume to be stable during the short-run, aggregate demand can be increased by increasing investment. Once investment increases, employment and income increase. Increased income leads to a rise in the demand for consumption goods which leads to further increase in employment and income. Once set in motion, employment and income tend to rise in a cumulative manner through the multiplier process till they reach the equilibrium level. According to Keynes, the equilibrium level of employment will be one of *under-employment equilibrium* because when income increases consumption also increases but by less than the increase in income. This behaviour of the consumption function widens the gap between income and consumption which ordinarily cannot by filled up due to the lack of required investment. The full employment income level can only be established if the volume of investment is increased to fill the income-consumption gap corresponding to full employment.

* For its understanding, the reader should so hold the figure that the word INCOME is towards him.

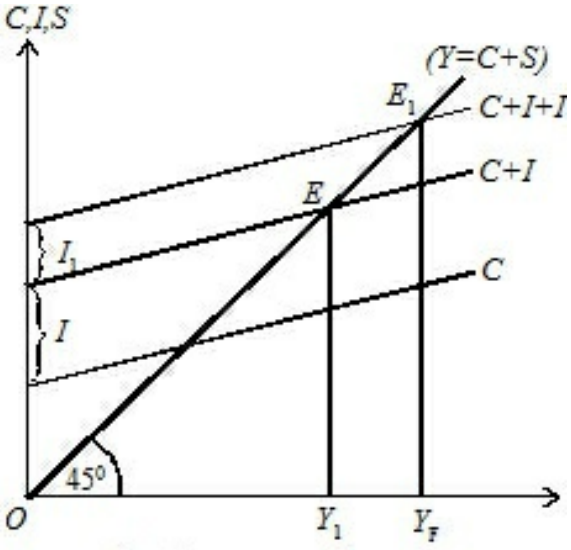


Fig. 2

The Keynesian cross model of under-employment equilibrium is explained in Figure 2 where income and employment are taken on the horizontal axis and consumption and investment on the vertical axis. Autonomous investment is taken as a first approximation. *C+I* is the aggregate demand curve plotted by adding to consumption function *C* an equal amount of investment at all levels of income. The 45° line is the aggregate supply curve. The economy is in equilibrium at point *E* where the aggregate demand curve *C+I* intersects the 45° line. This is the point of *effective demand* where the equilibrium level of income and employment OY_1 is determined. This is the level of *under-employment equilibrium* and not of full employment. There are no automatic forces that can make the two curves cross at a full employment income level. If it happens to be a full employment level, it will be accidental. Keynes regarded the under-employment equilibrium level as a normal case and the full employment income level as a special case.

Suppose OY_F is the full employment income level. To reach this level, autonomous investment is increased by I_1 so that the *C+I* curve shifts upward as $C+I+I_1$ curve. This is the new aggregate demand curve which intersects the 45° line (the aggregate supply curve) at E_1, the higher point of effective demand corresponding to the full employment income level OY_F . This also reveals that to get a desired increase in employment and income of Y_1Y_F, it is the multiplier effect of an increase in investment by I_1 ($=I_2$ in Panel C of Figure 1) which leads to an increase in employment and income by Y_1Y_F through successive rounds of investment.

SELECTED READINGS

J.M. Keynes, *General Theory of Employment, Interest and Money*, Ch. 18. 1936.

A.H. Hansen, *A Guide to Keynes*, Ch. 9. 1953.

G. Ackley, *Macroeconomic Theory*, Ch. XIII, 1961.

EXERCISES

1. How is equilibrium level of income determined under the Keynesian system ?
2. Is it possible to have equilibrium in the level of national income without achieving full employment ?
3. "The level of employment is determined by the level of income which in turn depends upon aggregate demand." Comment.
4. Explain Keynes's under-employment equilibrium.

CHAPTER 21

THE CLASSICAL VS. KEYNESIAN MODELS OF INCOME AND EMPLOYMENT

GENERAL THEORY: EVOLUTIONARY OR REVOLUTIONARY

The nineteen-thirties was the most turbulent decade that set off the most rapid advance in economic thought with the publication of Keynes's *General Theory of Employment, Interest and Money* in 1936. Keynes attacked the classical doctrine for its failure to solve the economic problems of the modern world. Around the turn of the present century, the world witnessed a series of crises which cast doubt on the practical utility of the orthodox economics. The Great Depression of the thirties demolished whatever faith was left of the self-regulating capitalist system. Thus the *General Theory* was born in a favourable environment and was characterised by economists like Harris as "The New Economics" and by others as revolutionary or evolutionary. But as pointed out by Professor Harris, "It is a matter of judgement whether the General Theory is simply classical economics further developed or embroidered, or whether Keynesian economics represents a genuine break."[1]

There has been a public debate in the academic journals among the economists on the occasion of the twentieth and twenty-fifth anniversaries of the publication of the *General Theory*, in fact right from its publication, as to whether it is evolutionary or revolutionary. No person is original in any pursuit of knowledge. He draws heavily from the ideas of the successive creative minds and formulates new ideas on their work and thought. Keynes also did the same thing. He accepted the classical theory, criticised and extended it and at the same time rejected parts of it. The

main elements of the *General Theory* can be found embryonic form in the works of his predecessors but Keynes's novelty lies in giving them a new complexion. As rightly observed by Harris, "Out of the straws of his predecessors, with some additions of his own, he had built a structure which no economist or economic practitioner can afford not to inspect or use." No doubt the Keynesian economics is built on the classical economics but it differs significantly from the latter in terms of assumptions, presentation of tools of analysis and policy measures. In this sense it is revolutionary rather than evolutionary. Keynes possessed great intuitive power and confidence for he wrote to *George Bernard Shaw* in 1935 before the publication of his *General Theory*, "You have to know that I believe myself to be writing a book on economic theory which will largely revolutionise nor, I suppose at once, but in the course of the next ten years—the way the world thinks about economic problem." Undoubtedly, the Keynesian analysis has significantly influenced matters of policy in the capitalist economics of the world. The following points mark Keynesian theory as revolutionary and a genuine departure from the classical economics.

---

1. S.E. Harris, *The New Economics: Keynes's Influence on Theory and Public Policy*, 1947, p. 55.

(1) Full Employment. The classicists believed in the existence of full employment in the economy and a situation of less than full employment was regarded, as abnormal. They, therefore, never thought it necessary to have a special theory of employment. On the other hand, Keynes considered the existence of full employment in the economy as a special case. He put forth a general theory of employment applicable to every capitalist economy. His notion of underemployment equilibrium is indeed revolutionary and has stood the test of the time.

(2) Say's Law. The classical analysis was based on Say's Law of Markets that "supply creates its own demand." The classicists thus ruled out the possibility of over production. "Keynes's greatest achievement," according to Prof. Sweezy "was the liberation of Anglo-American economics from this tyrannical dogma."[2] Keynes propounded the opposite view that *demand creates its own supply*. Unemployment results from the deficiency of effective demand because people do not spend the whole of

their income on consumption. Thus the development of the principles of effective demand and consumption function is a revolutionary contribution of Keynes to economic theory. To Klein, "The revolution was solely the development of a theory of effective demand,"[3] and to Hansen, "Consumption function is an epoch making contribution to the tools of economic analysis."[4]

(3) Laissez-Faire. The classical economics was based on the *laissez-faire* policy of a self-adjusting economic system with no government intervention. Keynes discarded the policy of *laissez-faire* because he believed that enlightened self-interest did not always operate in the public interest and it was this policy which led to the Great Depression. He, therefore, favoured state intervention and stressed the importance of public investment to fill the gap created by the deficiency of private investment. "Viewing Keynes's theory as a whole, its revolutionary nature lies," according to Prof. Dillard, "in the repudiation of any presumption in favour of laissez-faire."

---

2. *P.M. Sweezy,* op.cit., *p. 195.*

3. *L.R. Klein,* op.cit., *p. 56.*

4. *A.H. Hansen,* A Guide to Keynes, *1953.*

(4) Wage Cut. Pigou, one of the foremost classical economists, favoured the policy of wage-cut to solve the problem of unemployment. But Keynes opposed such a policy both from the theoretical and practical points of view. Theoretically, a wage-cut policy increases unemployment instead of removing it. Practically, workers are not prepared to accept a cut in money wage. Keynes, therefore, favoured a flexible monetary policy to a flexible wage policy to raise the level of employment in the economy. Prof. Harris regards Keynes's views on wages and employment as revolutionary.

(5) Saving. The classicists emphasized the importance of saving or thrift in capital formation for economic growth. To Keynes, saving was a private virtue and a public vice. Increase in aggregate saving leads to a decline in aggregate consumption and demand thereby decreasing the

level of employment in the economy. Keynes thus advocated public spending instead of public saving to remove unemployment. He thus 'smashed the last pillar of the bourgeois argument' that unequal income led to increased saving and to capital formation for growth. This view might be termed revolutionary.

(6) Saving-Investment Equality. The classicists believed that saving and investment were equal at the full employment level and in case of any divergence the equality was brought about by the mechanism of rate of interest. Keynes held that the level of saving depended upon the level of income and not on the rate of interest. Similarly investment is determined not only by rate of interest but by the marginal efficiency of capital.

(7) Trade Cycles. The classical economists failed to provide an adequate explanation of the cyclical phenomena. They could not explain the turning points of the business cycle satisfactorily and generally referred to boom and depression. Keynes' real contribution to the business cycle analysis lies in his explanation of turning points of the cycle and in the change of attitude as to what should and should not be done by the government to control the cycle. In this field, as opined by Mrs. Robinson, "Keynesian revolution commands the field."

(8) Monetary Theory. The classicists artificially separated the monetary theory from the value theory. Keynes, on the other hand, integrated monetary theory and value theory. He also brought interest theory into the domain of monetary theory. He regarded the rate of interest as a purely monetary phenomenon. He emphasized the demand for money as an asset and separated it into transactions demand, precautionary demand and speculative demand to explain the determination of the rate of interest in the short-run. By integrating the value theory and monetary theory through the theory of output, Keynes made money non-neutral as opposed to the classical view of neutrality of money.

(9) Macro Analysis. The classical economics was a microeconomic analysis which the orthodox economists tried to apply to the economy as a whole. Keynes, on the other hand, adopted the macro approach to economic problems. But the Keynesian revolution lies in its macro-dynamic orientation of aggregate income, employment, output,

consumption, demand, supply, saving and investment. As rightly pointed out by Prof. Hansen, "The *General Theory* has helped to make us think of economics in dynamic rather than in static terms."

---

5. *K. Kurihara (ed.),* Post Keynesian Economics, *1945.*

(10) Saving Capitalism. Keynes' most significant contribution lies in saving capitalism from the catastrophe it had fallen in the 1930's. The pure, unadulterated capitalism of the classical ideology could not function because as Keynes wrote, "It is not intelligent, it is not beautiful, it is not just, it is not virtuous and it does not deliver the goods." Keynes reformed capitalism by advocating the necessity of state intervention in order to increase aggregate demand and employment and thus saved it from giving way to communism. "And in this sense," observes Prof. Galbraith, "Keynes was pretty successful because it brought Marxism in the advanced countries to a halt."

(11) Policies. The classical economists being the votaries of *laisse-faire* policy had no faith either in fiscal policy or monetary policy. They believed in the balanced budget policy. Keynes, on the other hand, stressed the importance of deficit budgets during deflation and surplus budgets during inflation along with cheap money and dear money policies respectively. He was thus a practical economist whose models clarify both inflationary and deflationary episodes, and prosperous and depressed economies. His policy measures have been adopted by almost all the capitalist economies of the world. Thus in the words of Joan Robinson, "The Keynesian revolution has destroyed the old soporific doctrines *and* we are left in the uncomfortable situation of having to think for ourselves."

We may conclude that the *General Theory* is not evolutionary but is revolutionary in both economic thought and policy and is a genuine departure from the classical thought.

CRITICISMS OF KEYNESIAN THEORY

Despite the theoretical and practical significance of the Keynesian theory,

it is necessary to examine its failures and weaknesses for a proper evaluation. "Keynes while providing indispensable tools of analysis...raised more questions than he answered," according to Professor Kurihara.[5] Few would gainsay that Keynes "opened up new vistas and new pathways to a whole generation of economists," but many would now consider his analysis less than adequate for meeting such special problems as cyclical forecasts and controls, persistent inflation, the maintenance of full employment, booms, secular growth, non-linear structural relations and macro-functional distribution. These problems lie generally outside the scheme of the *General Theory*. Besides, every bit of Keynesian analysis has been criticised, such as aggregate demand, aggregate supply, the consumption function, the investment function, the monetary theory, etc. We study some of the major criticisms below.

(1) Aggregate Demand. Keynes asserted that the level of employment depended upon the level of aggregate demand which was in turn determined by the inactive consumption demand and active investment demand. And unemployment resulted from the lack of aggregate demand.According to Professor Schlesinger, the Keynesian theory of aggregate demand suffered from certain inherent defects which made his theory of employment unrealistic. He opines that "over-all demand is of course, to some extent, affected by relations on the supply side, Keynes's treatment of demand was therefore over-simple in that it neglected the possibility that the relative prices prevailing in the different sectors determine, in part, the total amount of outlays."

(2) Aggregate Supply. Professor Don Patinkin regards Keynes' treatment of the aggregate supply function inadequate. The aggregate supply is regarded as stable during the short-run. Moreover, the representation of the aggregate supply curve by the 45° line in the Keynesian cross diagram conveys the meaning that "demand creates its own supply." In other words, it implies that the aggregate supply is governed by aggregate demand. According to Patinkin, "This line of reasoning is yet another fallacious by-product of the usual Keynesian neglect of the supply side of the commodity market."

(3) Effective Demand. Economists have criticised Keynes's principle of

effective demand for two reasons. *First*, for taking the aggregate supply to be stable as noted above. *Second*, for assuming a direct functional relationship between effective demand and the volume of employment. According to Hazlitt, the volume of employment is not a function of effective demand, rather it depends upon the inter-relationship between wage-rates, prices and the supply of money. For instance, it is possible to achieve full employment even when the effective demand is low, provided wage-rates are so flexible that they could be adjusted quickly to the prices. Thus the direct relationship between effective demand and the volume of employment is fallacious. According to Prof. Burns, the determination of Keynes's theory in terms of effective demand "reflects a pleasant but dangerous illusion."

(4) Consumption Function. Keynes's consumption function though regarded as an epoch-making contribution to the tools of economic analysis yet it is not free from defects. "The relationship does not run simply from *current* income to *current* consumption," as Keynes forged, rather it "involves some complex average of past and expected income and consumption," as pointed out by Professor Ackley. According to Slichter, "The level of consumption is determined to a significant extent by conditions other than the level of real income which Keynes neglects altogether. They are the wealth effect, technological change, education, expectations, attitudes toward assets, etc.

(5) Investment Function. Keynes has also been criticised for formulating the functional relationship between investment and the rate of interest The influence of the rate of interest in determining the volume of investment is very uncertain. It was for this purpose that Keynes made his analysis more complicated by introducing the interrelation between the rate of interest and the marginal efficiency of capital to determine the level of investment.

Keynes erred in depending exclusively on the investment function and taking the consumption function to be stable in determining the volume of employment. It has been proved beyond doubt that raising the propensity to consume even during the short-run has a salutary effect on the volume of employment. Further, Keynes neglected the relationship between capital stock and investment.

Lastly, his theory of investment failed to consider the effect of investment on technological progress. According to Professor Slichter, "His theory of investment exaggerated the disposition to hoard and gratuitously assumed that the economy possesses only a meagre capacity to discover or to create investment opportunities." Thus Keynes ignores the impact of technology on the economy.

(6) Rate of Interest. The Keynesian theory of interest rate determination has been severely criticised by post-Keynesian economists. keynes made the rate of interest determined by the demand for and supply of money. The demand for money arises from the transactions motive, the precautionary motive and the speculative motive. Only the speculative demand for money is regarded interest elastic whereas the transactions demand is considered interest inelastic. According to Hansen, Keynes believed like the quantity theorists that the transactions demand for money was interest inelastic. But he was wrong because it is also interest elastic though at high interest rates.

Keynes's treatment of the speculative demand for money is very narrow because he confined himself only to cash and bonds, and failed to consider other types of assets. There is *"money illusion"* in the Keynesian speculative demand for money which means that the increased supply of money is absorbed only at a lower rate of interest.

Moreover, Keynes ignored what Patinkin calls the "direct influence of real-balance effect on aggregate demand." When the wealth of the people increases, it affects consumption and hence the demand for money.

Further, Keynes failed to consider the influence of price expectations on the demand for money. He assumed wages and prices to be given. Prof. Friedman in his *Restatement of the Quantity of Money* considers the demand for money as dependent on the rate of change in the level of prices among other factors. Under normal circumstances, the demand for money remains stable but during hyper-inflation the demand for money falls due to the effects of price-level expectations.

Lastly, Keynes has also been criticised by Harrod for formulating his theory in "stock" terms and neglecting the "flow" variables. This

weakness stems from his efforts to formulate a pure monetary theory of interest and his rejection of the Wicksellian natural rate of interest. Thus Keynes failed to incorporate the real forces determining the interest rate. As pointed out by Joan Robinson, "Keynes's theory treated the rate of interest as determined by the demand and supply of money. This was a useful simplification in the pioneering days of the theory...but there is no such thing as the rate of interest and that the demand and supply of every type of asset has just much right to be considered as the demand and supply of money."

(7) Expectations. Keynes has been criticised for his over emphasis on expectations. Expectations breed uncertainty. Though Keynes gave a dominant role to expectations in influencing the marginal efficiency of capital, yet he failed to formulate a precise theory of expectations. He relied on “convention” for forecasting changes in business expectations and failed “to confront *ex-ante* and *ex-post* reasoning," as Professor Hart puts it. "The essence of this convention," according to Keynes, "lies in assuming that the existing state of affairs will continue indefinitely, except in so far as we have specific reasons to expect a change." The reliance on the convention hypothesis makes Keynes's concept of expectations superfluous and unrealistic.

(8) Saving and Investment. Keynes did not pay as much importance to saving as to investment in his analysis. This stems from his weakness to relate saving as an *ex-post* factor pertaining to the current period. It is *ex-ante* saving that is more important in influencing the level of employment. Moreover, Keynes failed to recognise that saving is not hoarded but spent on both consumer and capital goods.

Another weakness of the Keynesian analysis pertains to the relation between saving and investment. On the one hand, Keynes regarded saving and investment as "merely diferent aspects of the same thing" and thus "necessarily equal." On the other, they were regarded as "two essentially different activities without even a nexus" so that they tended to equality only in equilibrium. Thus Keynes makes the saving-investment relationship very confusing.

(8) Wages. Economists have criticised the Keynesian analysis of wages

and employment. The Keynesian under-employment equilibrium is based on wage rigidity. Keynes also suggested increase in money wages or reduction of real wages to remove unemployment. Patinkin has shown that under-employment equilibrium "can exist even in a system of perfect competition and wage and price flexibility." Hazlitt holds that "the market mechanism applies to the labour market. When money wages are very high, there will be unemployment on the principle that when the price of any commodity is very high, the whole of it will not be sold." Patinkin's argument is more convincing "that a deficiency in commodity demand can generate a decrease in labour input without requiring a *priori* increase in real wage rate."

(10) Business Cycles. Keynes has also been criticised for his analysis of business cycles which was primarily based on expectations. Saulnier points out that Keynes *Notes on the Trade Cycle* lack in factual proof. In his words, "Keynes makes no attempt...to test any of his deductions with facts." Further, some of the variables of Keynes's business cycles such as expectations, marginal efficiency of capital and investment cannot explain the turning points of the business cycle. Keynes attributes the downturn to a sudden collapse in the marginal efficiency of capital. According to Hazlitt, the term marginal efficiency of capital being vague and ambiguous, "Keynes's explanation of the crisis of the marginal efficiency of capital is either a useless truism or an obvious error."

One of the serious omissions of Keynes's theory is the *acceleration principle*. This made his theory of business cycles one-sided because his explanation centres round the principle of multiplier. As pointed out by Hicks, "The theory of the acceleration and the theory of multiplier are the two sides of the theory of the fluctuations, just as the theory of demand and the theory of supply are the two sides of the theory of value."

(11) Dynamic Theory. Keynes regarded his theory dynamic and called it "the theory of shifting equilibrium." Even his foremost pupil Roy Harrod calls him "the father of dynamic economics." Keynes introduced an element of dynamism in his theory through the 'expectations.' But his analysis was concerned with the level of employment *at any time.* It is a lagless analysis. According to Prof. Kurihara, "The 'dynamic' nature of

Keynes' shifting equilibrium suggests that he is thinking dynamically, since there can be no shift from one position of equilibrium to another without prior movements of variables through time. Keynes made no attempt to show the process of transition from one position of equilibrium to another, however. His method of comparing different equilibrium levels of income has been termed comparative statics. Prof. Ackley calls the Keynesian model as "too static."

(12) Short-run Economics. Another criticism of the Keynesian economics is that it is applicable to the short-run. Keynes himself remarked, “In the long-run, we are all dead.” He, therefore, assumed a given stock of capital equipment, existing technique, tastes and habits of the people, organisation, size of population, etc. But all these factors change during the short-run. This makes Keynes’s analysis unrealistic. Moreover, economics is an incomplete study without concentrating on the long-run effects of these forces on the economy.

(13) Too Aggregative. The Keynesian model has been criticised for being “too aggregative.” In other words, it lays too much emphasis on the macro aspect and utterly neglects the micro aspect. The use of the aggregative concepts 'dooms the model to give wrong or misleading advice,' in the opinion of Professor Ackley. "The unit of analysis must be the individual commodity or commodities grouped in some other manner, e.g. by degree of elasticity of supply." For a real understanding of the working of macroeconomic variables like income, investment, consumption, employment, etc., the study of their microbehaviour is essential. Thus the aggregative nature of Keynesian economics detracts from its utility as a realistic study of economic problems.

(14) Closed Economy. The Keynesian theory is based on the assumption of a closed economy which excludes the impact of foreign trade on the level of employment and income. This makes Keynes’ analysis unrealistic because all economies are open economies, and foreign trade has an important impact on their level of employment. For instance, an unfavourable balance of trade leads to the flow of income abroad which results in the reduction of domestic income, investment and the volume of employment via the reverse operation of the multiplier. On the contrary, a

favourable balance of trade has the impact of increasing the level of income, investment and employment in the economy. Thus Keynes's neglect of the repercussions of foreign trade on the volume of employment is a serious defect in his theory.

(15) Perfect Competition. Another weakness of the Keynesian theory is that it is based on the unrealistic assumption of perfect competition. This makes his theory inapplicable to socialist or communist societies where the entire economy is regulated by the state. There is no cyclical unemployment in such economies. Hence the question of the applicability of Keynesian theory in them does not arise. As Prof. Harris remarked aptly, "If communism comes, Keynes will be as dead as Ricardo."

The Keynesian theory is not applicable even to modern capitalist economies where there is monopolistic competition rather than perfect competition. For instance, the principle of effective demand states that when the aggregate demand curve is over the aggregate supply curve, the entrepreneurs employ more labourers in expectation of earning larger profits till the point of effective demand is reached. But it is not essential that the entrepreneurs must employ more workers if there is imperfect competition in order to reach equilibrium level of employment. Thus the Keynesian theory is divorced from reality.

(16) General Theory. Keynes considered his theory as a “general theory.” But as is clear from the above points, it is not a general theory but a *special theory* which is applicable only under static conditions in a perfectly competitive closed economy. Moreover, it fails to solve the problems of underdeveloped countries. The tools and assumptions on which the Keynesian economics is built are incapable of bringing the development of such economies. Thus the Keynesian economics can by no means be termed as a general theory. Prof. Harris is more realistic when he says, "Those who seek universal truths applicable in all places and at all times, had better not waste their time on the *General Theory*."

(17) Problem of Unemployment. Keynes has been criticised for tackling only cyclical unemployment, and neglecting other types of unemployment to be found in capitalist economies. He did not offer any solution to frictional unemployment and technological unemployment. The problem

of technological unemployment had been ignored by Keynes because he failed to visualise the rapid technological discoveries which took place in the advanced capitalist countries. Thus the Keynesian economics is incomplete in solving the problem of unemployment.

(18) Policy Implications. The policy implications of Keynesian economics have also been criticised. Some of the criticisms are discussed below.

(a) To fight unemployment, Keynes recommended the policy of deficit spending. But this policy has serious repercussions, because the state may spend beyond its means in an extravagant manner. Moreover, in America deficit spending by the government led to inflation instead of raising the volume of employment. As Prof. Hazlitt opines, "Inflation is at once an uncertain remedy for unemployment and unnecessary remedy for unemployment. To try to cure unemployment by inflation is to adjust the piano to the stool rather than the stool to the piano." Therefore, inflation or deficit spending cannot be relied upon to cure unemployment.

(b) Keynes's favoured public investment to overcome depression and to attain full employment. Though he said that the role of public investment was to supplement private investment and not to supplant it, yet public investment has tended to replace private investment to a large extent. With the nationalisation of road, air and rail transport, and a number of other industries, and the starting of state enterprises, the sphere of public sector has expanded considerably. This has narrowed the sphere of private enterprise.

(c) Keynes advocated progressive taxation to control inflationary trends in the economy. But higher taxes on companies may discourage private investment, and high commodity taxes may discourage consumption. This may have cumulative adverse effect on private investment, and thus lead the economy towards recession.

(d) Keynes paid little attention to monetary policy. In the Keynesian system money in neutral in situations of full employment and liquidity trap (when the rate of interest becomes inelastic in a depression). It is only in the intermediate situation between these two extremes that money is

non-neutral. This is a great weakness in the Keynesian analysis because monetary policy plays an important role even during these extreme situations, as has been proved by Friedman, Metzler, Patinkin and others.

(e) Keynes' policy measures fail to tackle the problems of capital formation and growth which result from technological innovations. They are also incapable of solving the problems of underdeveloped countries. In fact, the application of Keynesian policy measures to such economies have created more problems instead of solving them.

(f) Lastly, the Keynesian economics fails to provide solutions to a number of socio-economic problems facing the developed countries. Such problems include fair employment, income distribution and resource allocation. This is a serious weakness in Keynesian policy measures.

Conclusion

The critical evaluation of Keynesian economics reveals that there are the Keynesians who eulogise Keynes and there are the anti-Keynesians like Hazlitt who “could not find a single doctrine that was both true and original.” On the other hand, Keynes’ greatest follower Dillard writes, “Keynes was an original thinker in the sense that he arrived at his ideas in his own way. The ideas he advanced were his own even though some one else may have expounded the same or similar ideas at an early date.” Though the problems of today are somewhat different from what they were when Keynes wrote his *General Theory*, yet most economists approach the present day problems within the framework of the Keynesian analysis. Despite Samuelson's severe denunciation of the *General Theory* as a "badly-written book, poorly organised...not well-suited for classroom use, ...arrogant, bad tempered, polemical, not overly generous in its acknowledgements and *abounding* in meads and confusions," it still remains the most popular treatise on economics whose technical apparatus has been absorbed into the general body of economics. There is hardly any book on macroeconomics, monetary economics and public economics which is without the imprint of Keynesian thought and policy. Prof. Harry Johnson wrote in 1961, "At this date there is no need to labour the point that the *General Theory* deserves much of the credit for the fact that the maintenance of high and stable employment is now accepted as a

governmental responsibility, or that Keynes's theory of effective demand is the origin of the modern theory of economic policy." And according to Dillard, "The acceptance of deficit financing as a respectable type of public policy is one of the remarkable changes in public thinking for which Keynesian economics has been primarily responsible." We therefore, do not agree with Hazlitt, the staunch anti-Keynesian that the *General theory* was "one of the great intellectual scandals of our age." In fact, Schumpeter's assessment of Malthus applies fairly to Keynes. Keynes "had the good forture—for this is good fortune—to be the subject of equally unreasonable, contradictory appraisals. He was a benefactor of humanity. He was a fiend. He was a profound thinker. He was a dunce. The man whose work stirred people's minds so as to elicit such passionate appraisals was *ipso facto* no mediocrity." Rather, he was a genius.

SELECTED READINGS

Robert Lekachman (Ed.). *Keynes' General Theory: Roport of Three Decades,* 1964.

N. Marshall (Ed.), *Keynes: Updated or Outdated?*, 1970.

S.E. Harris (Ed.), *The New Economics: Keynes' Influence on Theory and Public Policy,* 1947.

EXERCISES

1. 'The General Theory is simply Classical Economics further developed or embroidered by Keynes." Discuss.
2. Evaluate the Keynesian theory of income and employment.
3. Discuss the theoretical and practical significance of Keynesian theory of employment.
4. Discuss Keynes' influence on economic thought and policy.
5. Compare the classical and Keynesian models of income determination and point out the fundamental differences between the two.

CHAPTER 22

UNEMPLOYMENT AND FULL EMPLOYMENT

INTRODUCTION

Unemployment has been one of the most persistent and unmanageable problems facing all industrial countries of the world. At the same time, the goal of public policy has been to remove unemployment and to achieve full employment in such countries. We attempt below the various types of causes of unemployment for an understanding of the meaning of the term full employemnt.

TYPES OF UNEMPLOYMENT

Before explaining the various types of unemployment, it is necessary to define the term *unemployment*. Everyman's Dictionary of Economics defines unemployment as "involuntary idleness of a person willing to work at the prevailing rate of pay but unable to find it." It implies that only those persons are to be regarded as unemployed who are prepared to work at the prevailing rate of pay but they do not find work. Voluntarily unemployed persons who do not want to work like the idle rich, are not considered unemployed. We now analyse the various causes or types of unemployment.

1. Frictional Unemployment. Frictional unemployment exists when there is lack of adjustment between demand for and supply of labour. This may be due to lack of knowledge on the part of employers about the availability of workers or on the part of workers that employment is available at a particular place. It is also caused by lack of necessary skills

for a particular job, labour immobility, breakdowns of machinery, shortages of raw materials, etc. The period of unemployment between losing one job and finding another is also included under frictional unemployment.

2. Seasonal Unemployment. Seasonal unemployment results from seasonal fluctuations in demand. Employment in ice factories is only for the summer. Similarly ice-cream sellers remain unemployed during winter and chestnut-sellers during summer. The same is the case with agricultural workers who remain employed during harvesting and sowing seasons and remain idle for the rest of the year.

3. Cyclical Unemployment. Cyclical unemployment arises due to cyclical fluctuations in the economy. They may also be generated by international forces. A business cycle consists of alternating periods of booms and depressions. It is during the downswing of the business cycle that income and output fall leading to widespread unemployment.

4. Structural Unemployment. Structural unemployment results from a variety of causes. It may be due to lack of the co-operant factors of production, or chages in the economic structure of the society. The word structural implies that "the economic changes are massive, extensive, deep-seated, amounting to transformation of an economic structure, *i.e*., the production functions or labour supply distribution. More specifically, it refers to changes which are large in the particular area, industry or occupation." Shifting patterns in the demand for the products of various industries have also been responsible for this type of unemployment. There are, however, economists who argue that the higher unemployment in America since 1957 has been due to causes other than inadequae demand: (1) A faster rate of technological change; (2) a displaced worker remains unemployed for a number of days in finding a new job; and (3) most of the unemployed workers belong to blue-collar groups. The supporters of the structural transformation thesis hold that the number of vacancies is greater than or equal to the number of displaced workers due to structural changes in a particular area, industry or occupation, and that unemployment is not due to inadequacy of demand.

5. Technological Unemployment. Keynes failed to take into account

technological unemployment that has taken place more rapidly in the post-war period. Modern prodution process is essentially dynamic where innovations lead to the adoption of new machineries and inventions thereby displacing existing workers leaving behind a trail of unemployment. When there is automation or displacement of old technology by a new one requiring less workers than before, there is technological unemployment. A special case of technological unemployment is that "which is not due to improvements in the technique of production but in the technique of *organisation*." It pertains to making management more efficient which may decide upon modernising existing facilities or closing down obsolete plants. In all such cases unemployment is bound to decrease.

In fact, there is little to distinguish between *structural* and *technology* unemployment. One of the causes of structural unemployment is technological change. Technological change itself causes obsolesence of skills thereby leading to structural unemployment. Further, both structural and technological unemployment are related to inadequate demand. Technological change tends to increase output per man-hour which has the effect of raising the potential total output in the economy. If this *potential* growth in output is not matched by the *actual* growth in output, there will be unemployment in the economy due to deficiency in demand. Therefore, modern economists are of the view that unemployment is caused by structural changes, technological changes and by inadequacy of demand taken together.

6. Disguised Unemployment. Disguised or concealed unemployment or underemployment is a notable feature of underdeveloped countries. Such unemployment is not voluntary but involuntary. People are prepared to work but they are unable to find work throughout the year due to the lack of complementary factors. Such unemployment is found among rural landless and small farmers due to the seasonal nature of farm operations and inefficient land and equipment to keep them fully employed. A person is said to be disguised unemployed if his contribution to output is less than what he can produce by working for normal hours per day. His marginal productivity is nil or negligible, and by withdrawing such labourers, farm output can be increased.

There are also other types of underemployed persons in such countries. A person is considered to be underemployed if he is forced by unemployment to take a job that he thinks is not adequate for his purpose, or not commensurate with his training. Further, there are those who work full time in terms of hours per day but earn very little to rise above the poverty level. They are hawkers, petty traders, rickshaw pullers, workers in hotels and restaurants and in repair shops, etc. in urban areas. Open and disguised unemployed in urban and rural areas are estimated at 30-35 per cent of the labour force in underdeveloped countries.

MEANING OF FULL EMPLOYMENT

Right from the classicals to the modern economists, there is no unanimity of views on the meaning of 'full employment'. It is a very "slippery concept", according to Professor Ackley. But the credit for popularising it goes to Keynes, and since the end of the Second World War it has been accepted as one of the important goals of public policy. Though "full employment is not definable nor should it be defined," according to Professor Henry Hazlitt, yet it is worth-while analysing the various views of economists on full employment.

The Classical View[1]

The classical economists always believed in the existence of full employment in the economy. To them full employment was a normal situation and any deviation from this was regarded as something abnormal. According to Pigou, the tendency of the economic system was to automatically provide full employment in the labour market. Unemployment resulted from the rigidity in the wage structure and interference in the working of free market system in the form of trade union legislation, minimum wage legislation, etc. Full employment exists "when everybody who at the running rate of wages wishes to be employed." Those who are not prepared to work at the existing wage rate are not unemployed in the Pigovain sense because they are voluntarily unemployed. There is, however, no possibility of involuntary unemployment in the sense that people are prepared to work but they do not find work. According to Pigou, "With perfectly free competition—

there will always be at work a strong tendency for wage rates to be so related to demand that everybody is employed." However, this classical view on full employment is consistent with some amount of frictional, voluntary, seasonal or structural unemployment.

---

1. For figures and detailed explanation refer to Ch. 6.

The Keynesian View

According to Keynes, full employment means the absence of involuntary unemployment. In other words, full employment is a situation in which everybody who wants to work gets work. Full employment so defined is consistent with frictional and voluntary unemployment. Keynes assumes that "with a given organisation, equipment and technique, real wages and the volume of output (and hence of employment) are uniquely co-related, so that, in general, an increase in employment can only occur to the accompaniment of a decline in the rate of wages." To achieve full employment, Keynes advocates increase in effective demand to bring about reduction in real wages. Thus the problem of full employment is one of maintaining adequate effective demand. "When effective demand is deficient," writes Keynes, "there is underemployment of labour in the sense that there are men unemployed who would be willing to work at less than existing real wage. Consequently, as effective demand increases, employment increases, though at a real wage equal to, or less than, the existing one, until a point comes, at which there is no surplus of labour available at the then existing real wage." Keynes gives an alternative definition of full employment at another place in his *General Theory* thus: "It is a situation in which aggregate employment is inelastic in response to an increase in the effective demand for its output." It means that the test of full employment is when any futher increase in effective demand is not accompained by any increase in output. Since the supply of output becomes inelastic at the full employment level, any further increase in effective demand will lead to inflation in the economy. Thus the Keynesian concept of employment involves three conditions: (i) reduction in the real wage rate, (ii) increase in effective demand, and (iii) inelastic supply of output at the level of full employment.[2]

Other Views on Full Employment

According to Professor W.W. Hart, attempting, to define full employment raises many people's blood pressure. Rightly so, because there is hardly any economist who does not define it in his own way.

Lord Beveridge in his book *Full Employment in a Free Society* defined it as a situation where there were more vacant jobs than employed men so that normal lag between losing one job and finding another will be very short. By full employment he does not mean zero employment which means that full employment is not always full. There is always a certain amount of frictional unemployment in the economy even when there is full employment. He estimated frictional unemployment of 3% in a full employment situation for England. But his pleading for more vacant jobs than the unemployed cannot be accepted as the full employment level.

According to the American Economic Association Committee, "Full employment means that qualified poeple who seek jobs at prevailing rates can find them in productive activities without considerable delay. It means full time jobs for people who want to work full time. It does not mean people like house-wives and students are under pressure to take jobs when they don't want jobs or that workers are under pressure to put in undesired overtime. It does not means unemployment is ever zero" This is not a definition but a description of full employment situation where all qualified persons who want jobs at current wage rates find full-time jobs. Here again, like Beveridge, the Committee considered full employment to be consistent with some amount of unemployment.

---

2. See Figure 2 of Ch. 20 and its explanation.

Individual economists may, however, continue to differ over the definition of full employment, but the majority has veered round the view expressed by the U.N. Experts on *National and International Measures for Full Employment* that "full employment may be considered as a situation in which employment cannot be increased by an increase in effective demand and unemployment does not exceed the minimum allowances that must be made for the effects of frictional and seasonal factors." This

definition is in keeping with the Keynesian and Beveridgian views on full employment. It is now agreed that full employment stands for 96 to 97 per cent employment, with 3 to 4 per cent unemployment existing in the economy due to frictional factors.

Measures to Achieve and Maintain full Employment

Since underemployment is caused by deficiency in effective demand, full employment can be achieved by increasing effective demand either by stimulating investment or consumption, or both. Full employment is thus sought to be achieved and maintained by monetary, fiscal and direct measures which are discussed in the chapters on Monetary and Fiscal Policies.

Exercises

1. What do you mean by full employment ? Explain measures to achieve full employment.
2. Distinguish between structural unemployment and technological unemployment.
3. Write notes on : frictional, cyclical and disguised unemployment.

CHAPTER 23

APPLICABILITY OF KEYNES'S THEORY TO UNDERDEVELOPED COUNTRIES

INTRODUCTION

The Keynesian theory is not applicable to every socio-economic set-up. It only applies to advanced democratic capitalist economies. As Schumpeter wrote, "Practical Keynesianism is a seedling which cannot be transplanted into foreign soil; it dies there and becomes poisonous before it dies. But left in English soil this seedling is a healthy thing and promises both fruit and shade. All this applies to every bit of advice that Keynes ever offered."[1] Before we study the applicability of Keynesian economics to underdeveloped countries, it is essential to analyse the assumptions of Keynesian economics vis-a-vis the conditions prevailing in underdeveloped economies.

1. J.A Schumpter, *Ten Great Economists,* op. cit., p. 275.

KEYNESIAN ASSUMPTIONS AND UNDERDEVELOPED COUNTRIES

The Keynesian economics is based on the following assumptions which limit its applicability to underdeveloped countries.

(1) The Keynesian theory is based on the existence of *cyclical un-employment* which occurs during a depression. It is caused by deficiency in effective demand. Unemployment can be removed by an increase in the level of effective demand. But the nature of unemployment in an

underdeveloped country is quite different from that in a developed economy. In such economies unemployment is *chronic* rather than cyclical. It is not due to lack of effective demand but is the result of deficiency in capital resources. Apart from chronic unemployment, underdeveloped countries suffer from disguised unemployment. Keynes was concerned with the removal of involuntary unemployment and the problem of economic instability. So he did not refer to disguised unemployment and its solution. The remedy for the chronic and disguised unemployment is economic development to which Keynes paid no attention at all. Thus the Keynesian assumptions of cyclical unemployment and economic instability are hardly tenable in an underdeveloped economy.

(2) The Keynesian economics is a *short period anaylsis* in which Keynes takes "as given the existing skill and quantity of available labour, the existing quantity and quality of available equipment, the existing technique, the degree of competition, the tastes and habits of the consumer, the disutility of different intensities of labour and of the activities of supervision and organisation, as well as social structure."[2] The development economics, however, is a long period analysis in which all the basic factors assumed by Keynes as given, change over time.

(3) The Keynesian theory is based on the assumption of *closed economy*. But underdeveloped countries are not closed economies. They are open economies in which foreign trade plays a dominant role in developing them. Such economies primarily depend on the exports of agricultural and industrial raw materials and the imports of capital goods. Thus the Keynesian economics has little relevance to underdeveloped countries in this respect.

(4) The Keynesian Theory assumes an *excess supply of labour* and other *complementary resources* in the economy. This analysis refers to a depression economy where "the industries, machines, managers and workers, as well as consumption habits, are all there, only waiting to resume their temporarily suspended functions and roles."[3] But in underdeveloped economies, there is no temporary suspension of economic activity. Economic activity is static, capital, skills, factor supplies and

economic infrastructure are woefully lacking.

(5) Moreover, it can be inferred from the above assumption that *labour and capital are unemployed simultaneously*, according to the Keynesian analysis. When labour is unemployed, capital and equipment are also not fully utilised or there is excess capacity in them. But this is not so in underdeveloped countries. When labour is unemployed, there is no question of capital being unutilized because there is acute shortage of capital and equipment.

---

2. J.M. Keynes, *op. cit.*, p. 245 note 1.

3. A.O. Hirschman, *The Strategy of Economic Development*, p. 54.

THE KEYNESIAN TOOLS AND UNDERDEVELOPED COUNTRIES

Thus the assumption on which the Keynesian theory is based are not applicable to the conditions prevailing in underdeveloped countries. We now study the principal *tools* of the Keynesian theory to test their validity to underdeveloped countries.

1. Effective Demand. Unemployment is caused by the deficiency of effective demand, and to get over it, Keynes suggested the stepping up of consumption and non-consumption expenditures. In an underdeveloped country, however, there is no involuntary unemployment but disguised unemployment. Unemployment is caused not by lack of complementary resources. The concept of effective demand is applicable to those economies where unemployment is due to excess savings. In such a situation the remedy lies in stepping up the levels of consumption and investment through various monetary and fiscal measures. But in an underdeveloped economy income levels are extremely low, the propensity to consume is very high and savings are almost nil. All efforts to increase money incomes through monetary and fiscal measures will, in the absence of complemenatary resources, lead to price inflation. Here the problem is not one of raising the effective demand but one of raising the levels of employment and per capita income in the context of economic development. “The economic progress consists of two distinct categories:

one, where at the given level of economic development, you move from low employment to full employment, and the other, where you move from full employment at a given level of economic development to full employment at the next higher level of economic development. The Keynesian thesis applies only to *the first category*."[4]

2. Propensity to Consume. One of the important tools of Keynesian economics is the propensity to consume which highlights the relationship between consumption and income. When income increases, consumption also increases but by less than the increment in income. This behaviour of consumption further explains the rise in saving as income increases. In underdeveloped countries, these relationships between income, consumption and saving do not hold. People are very poor and when their income increases, they spend more on consumption goods because their tendency is to meet their unfulfilled wants. The marginal propensity to cousume is very high in such countries, whereas the marginal propensity to save is very low. The Keynesian economics tells us that when the *MPC* is high, the consumer demand, output and employment increase at a faster rate with the increase in income. But in an underdeveloped country, it is not possible to increase the production of consumer goods due to the scarcity of co-operant factors, when consumption increases with the rise in income. As a result, prices rise instead of a rise in the level of employment.

3. Saving. On the saving side, Keynes regarded saving as a social vice for it is excess of saving that leads to a decline in aggregate demand. Again, this idea is not applicable to underdeveloped countries because saving is the panacea for their economic backwardness. Capital formation is the key to economic development and capital formation is possible through increased saving on the part of people. Underdeveloped countries can progress by curtailing consumption and increasing saving, as opposed to the Keynesian view of raising consumption and reducing saving. To underdeveloped countries, saving is a virtue and not a vice.

---

4. V.K.R.V. Rao, *Essays in Economic Development*, op. cit., p. 61. Italics mine.

4. Marginal Efficiency of Capital. According to Keynes, one of the

important determinants of investment is the marginal efficiency of capital. There is an inverse relationship between investment and *MEC*. When investment increase, the *MEC* falls, and when investment declines, the *MEC* rises. This relationship is, however, not applicable to underdeveloped countries. In such economies, investment is at a low level and the *MEC* is also low. This paradox is due to the lack of capital and other resources, small size of the market, low demand, high costs, underdeveloped capital and money markets, uncertainties, etc. All these factors keep the *MEC* (profit expectations) and investment at a low level.

5. Rate of Interest. The rate of interest is the second determinant of investment in the Keynesian system. It is, in turn, determined by liquidity preference and the supply of money. Of the motives for liquidity preference, the transactions and precautionary motives are income elastic and they do not influence the rate of interest. It is only the demand for money for the speculative motive that affects the rate of interest. In underdeveloped countries, the liquidity preference for transactions and precautionary motives is high and for the speculative motive low. Therefore, liquidity preference fails to influence the rate of interest. The other determinant of the interest rate is the supply of money. According to Keynes, increase in the supply of money lowers the interest rate and encourages investment, income and the level of employment. But in underdeveloped countries, an increase in the supply of money leads to the rise in prices rather than to the fall in interest rate. As Keynes himself observed in citing India's example, "The history of India at all times has provided an example of a country impoverished by a preference for liquidity amounting to so strong a passion that even an enormous and chornic influx of the precious metals has been insufficient to bring down the rate of interest to a level which was compatible with the growth of real wealth."[5] Thus the rate of interest in underdeveloped countries is not influenced so much by the demand for and the supply of money as by traditions, customs and insitutional factors.

6. The Multiplier. Dr. V. K. R.V. Rao has analysed the feasibility of applying the Keynesian multiplier theory and policy implications to an underdeveloped country like India.[6] According to Dr. Rao, Keynes never formulated the economic problems of underdeveloped countries nor did he

discuss the relevance to these countries for either the objective or the policy that he proposed for the more developed countries. The result has been a rather unintelligent application of Keynesian economics to the problems of underdeveloped countries. The Keynesian concept of multiplier is based on the following *four assumptions*: (a) Involuntary unemployment, (b) an industrialized economy where the supply curve of output slopes upward to the right but does not become vertical till after a *substantial interval*, (c) excess capacity in the consumption goods industries, and (d) comparatively elastc supply of the working capital required for increased output.

Given these assumptions, if we apply the multiplier theory on underdeveloped countries, the value of the multiplier will be apparently much higher than even in a developed country. We know that the mutliplier depends on the size of the marginal propensity to consume. Since in an underdeveloped country the marginal propensity to consume is fairly high, small increments of investment are likely to induce full employment much earlier than in a rich country where the marginal propensity to consume is low. This is something paradoxical and contrary to facts. For the assumptions on which the multiplier theory is based do not hold valid in the case of an underdeveloped country. Let us test them in the light of conditions prevailing in an underdeveloped country like India.

---

5. J.M. Keynes, *General Theory*, p. 337.

6. V.K.R.V. Rao, *op. cit.*

(a) Involuntary unemployment in the Keynesian analysis is associated with a capitalist economy where the majority of workers work for wages and where production is more for exchange than for self-consumption. According to Prof. Das Gupta, the organized sector of an underdeveloped economy with its large-scale industries and fairly well-developed banking system comes under the scope of Keynesian economics, for it presents the features of a capitalistic economy. But involuntary unemployment in this sector is insignificant when considered in relation to the total working population of the country. As per Prof. Das Gupta's rough estimate, involuntary unemployment in India works out to be 0.2 per cent of the

total working force, on the assumption that 10 per cent of the persons employed in organized industry are involuntarily unemployed and that hardly 2 per cent of the total working population is absorbed by the organised industry.[7]

In fact, in an overpopulated underdeveloped country, there exists disguised unemployment. Ostensibly people are engaged in agriculture but if some of them are withdrawn from the farm, there will be no reduction in output. In an underdeveloped economy, the existence of disguised unemployment instead of involuntary unemployment hinders the working of the mutiplier theory. The secondary, tertiary and other effects of the initial increment do not follow mainly because there is no labour force willing to accept employment at the current wage level. The disguised unemployment are not available at the current wage level because, *firstly*, they are not conscious of the fact that they ae unemployed, and *secondly*, they are already receiving a real income which gives them at least as much satisfaction as they would get from the current wage level. Thus the absence of involuntary unemployed and the presence of disguised unemployment in underdeveloped countries retard the operation of the multiplier toward increasing output and employment.

(b) The supply curve of output in an underdeveloped country is inelastic which renders the working of the multiplier all the more difficult. The reason is that the nature of the consumption goods industries is such that they are unable to expand output and offer more employment. The main consumption goods industry in an underdeveloped country is agriculture which is almost stagnant. The supply curve of agricultural output is backward sloping so that an increase in the value of output does not necessarily lead to an increase in the volume of output. This is because in the short-run necessary facilities are not available to the agricultural producers for increasing output. As a result, the secondary, tertiary and other increases in income, output and employment do no come about with an initial increment of investment. The primary increase in income is spent on food and its multiplier effect is lost.

(c) Since the marginal propensity to consume is high in underdeveloped countries, the inceased income is spent on self consumption of food

products by the farmers which leads to a reduction of the marketable surplus of food grains. This, in turn, leads to a rise in the prices of foodgrains in the non-agricultural sector without a rise in the aggregate real income. The possibility of spending more by the agriculturists on non-agricultural goods is, however, limited because there is little excess capacity in industries.

---

7. A. K. Dass Gupta, *Planning and Economic Growth,* pp. 32-33.

(d) Output is difficult to increase due to the non-availability of sufficient raw materials, capital equipment and skilled labour. Thus, concludes Dr. Rao, "the primary increase in investment, and, therefore, increase in income and employment leads to a secondary and a tertiary increase in income, but not any noticeable increase in output or employment, either in the agricultural or in the non-agricultural sector."[8]

Thus the absence of conditions (c) and (d) in an underdeveloped country renders the operation of the multiplier difficult.

Conclusion. The obvious conclusion is that the Keynesian principle of multiplier does not operate in an underdeveloped country like India mainly due to two reasons: *firstly*, involuntary unemployment of the Keynesian type is not to be found, and *secondly*, the supply of agricultural and non-agricultural output is inelastic due to the working of certain factors peculiar to such economies.

7. Policy Measures. Not only this, even the Keynesian policy prescriptions are hardly tenable under the conditions prevailing in underdeveloped countries. Dr. Rao maintains that an attempt to increase investment through deficit financing leads to an inflationary rise in prices rather than to an increase in output and employment. He is, therefore, of the view that "the economic policy of deficit financing and disregard for thrift advocated by Keynes for securing full employment does not apply in the case of an underdeveloped country." But in another essay,[9] he contends that deficit financing for capital formation does not lead to inflation since it is used for increasing the capacity and thereby imparting elasticity to the supply curve of output. However, a certain measure of

price rise is inevitable but it is of a "self-liquidating character." He points out that the history of war finance shows that forced savings obtained through a rise in prices, when used for capital formation, are nothing but deficit financing for economic development. "The only question is the extent in which it is wise to resort to deficit financing; and the obvious answer is that deficit financing should not be restorted to beyond the point at which it becomes inflationary."

Prof. Das Gupta advocates the use of Keynesian policy of public investment to achieve a higher standard of living and to provide increasing employment opportunities in underdeveloped countries. But in the absence of adequate public savings and the flow of foreign capital, he advocates deficit financing which if not accompanied by a system of price and capital issue controls, in the transitional period, will lead to inflationary rise in prices. He, however, emphasises that "we would be deceiving ourselves if we thought that the process could function within the framework of a liberal economy, such as Keynes sought so anxiously to preserve."[10] For underdeveloped countries "the old fashioned prescription of work harder and save more still seems to hold as the medicine for economic progress"[11] than the Keynesian hypothesis that consumption and investment should be increased simultaneously. But it cannot be denied that though the Keynesian policy prescriptions do not apply *in toto* to the problems of underdeveloped countries, yet the Keynesian tools of analysis are indispensable for understanding the problems of such economies.

---

8. V.K.R.V. Rao, *op. cit.*, p. 41.

9. Ibid., Chapter 5.

10. A.K. Das Gupta, *op, cit*., p. 37.

11. V.K.R.V. Rao, *op, cit.*

To conclude with Prof. Das Gupta :"Whatever the generality of the *General Theory* may be in the sense in which the term 'general' was used by Keynes, the applicability of the propositions of the *General Theory* to conditions of an underdeveloped economy is at best limited."[12]

SELECTED READINGS

A. K.Das Gupta, Planning and Economic Growth, *Ch.2.*

V.K.R.V. Rao, Essays in Economic Development, *Ch. 2.*

EXERCISES

1. What assumptions has Keynes made in his *General Theory*? To what extent do you think the conditions in an underdeveloped economy affirm or refuse those assumptions?
2. Critically discuss the Keynesian theory of employment and show the extent to which its main conclusions are applicable in the case of economically underdeveloped countries.

---

12. A.K. Das Gupta. *op. cit.*, p.34.

PART-IV

MONETARY THEORY

CHAPTER 24

MONEY

NATURE AND DEFINITION OF MONEY

There has been lot of controversy and confusion over the meaning and nature of money. As pointed out by Scitovsky, "Money is a difficult concept to define, partly because it fulfils not one but three functions, each of them providing a criterion of moneyness ... those of a unit of account, a medium of exchange, and a store of value."[1] Though Scitovsky points toward the difficulty of defining money due to moneyness[2], yet he gives a wide definition of money. Coulborn defines money as "the means of valuation and of payment; as both the unit of account and the generally acceptable medium of exchange." Coulborn's definition is very wide. He includes in it the 'concrete' money such as gold, cheques, coins, currency notes, bank draft, etc. and also abstract money which "is the vehicle of our thoughts of value, price and worth." Such wide definitions have led Sir John Hicks to say that "money is defined by its functions: anything is money which is used as money: 'money is what money does.'[3] These are the functional definitions of money because they define money in terms of the functions it performs.

---

1. T. Scitovsky, *Money and the Balance of Payments,* 1969.

2. Moneyness means liquidity. All things which possess liquidity have moneyness. For its detailed study, refer to the chapter "The supply of Money.".

Some economists define money in legal terms saying that "anything which the state declares as money is money." Such money possesses general acceptability and has the legal power to discharge debts. But

people may not accept legal money by refusing to sell goods and services against the payment of legal tender money. On the other hand, they may accept some other things as money winch are not legally defined as money in discharge of debts which may circulate freely. Such things are cheques and notes issued by commercial banks. Thus besides legality, there are other determinants which go to make a thing to serve as money.

THEORETICAL AND EMPIRICAL DEFINITIONS OF MONEY

There being no unanimity over the definition of money. Prof. Johnson[4] distinguishes four main schools of thought in this regard which are discussed below alongwith the views of Pesek and Saving.

1. The Traditional Definition of Money. According to the traditional view, also known as the view of the Currency School, money is defined as currency and demand deposits, and its most important function is to act as a medium of exchange. Keynes in his *General Theory* followed the traditional view and defined money as currency and demand deposits. Hicks in his *Critical Essays in Monetary Theory* points toward a threefold traditional classification of the nature of money: "to act as a unit of account (or measure of value as Wicksell put it), as a means of payment, and as a store of value." The Banking School criticised the traditional definition of money as arbitrary. This view about the meaning of money is very narrow because there are other assets which are equally acceptable as media of exchange. These include time deposits of commercial banks, commercial bills of exchange, etc. By ignoring these assets the traditional view is not in a position to analyse their influence in increasing their velocity. Further, by excluding them from the definition of money, the Keynesians place greater emphasis on the interest elasticity of the demand function for money. Empirically, they forged a link between the stock of money and output via the rate of interest.

2. Friedman's Definition of Money. The monetarist (or Chicago) view is associated with Friedman and his followers at the University of Chicago. By money Friedman means "literally the number of dollars people are carrying around in their pockets, the number of dollars they have to their credit at banks in the form of demand deposits and commercial bank time

deposits".[5] Thus he defines money as "the sum of currency plus all adjusted deposits in commercial banks".

This is the "working definition" of money which Friedman and Schwartz used for the empirical study of the monetary trends of the US for selected year 1929, 1935, 1950, 1955 and 1960. This was a narrower definition of money and the adjustment in both demand and time deposits of commercial banks was devised to take into account the increasing financial sophistication of the commercial banks and the community. But Friedman could not establish a single index of this sophistication. Even with this adjustment, cash and deposit monies were not strictly comparable over long periods.

---

3. John Hicks, *Critical Essay in Monetary Theory*, 1967.

4. H.G. Johnson, "Monetary Theory and Policy", *A.E.R.*, Vol, 52, June 1962.

5. M. Friedman, *Employment, Growth and Price level*, 1959.

However, the correlation evidence for 1950, 1955 and 1960 suggested a broader definition of money as "any asset capable of serving as a temporary abode of purchasing power". So Friedman gives two types of definitions of money. One on theoretical basis and the other on empirical basis. This led to a lot of controversy which Friedman tried to solve on the basis of methodological issues. According to him, "The definition of money is to be sought for not on grounds of principle but on grounds of usefulness in organising our knowledge of economic relationships." Thus the definition used for empirical purposes is unimportant because different definitions will give different results. The empirical results will ultimately depend upon the nature of assets included in the definition of money as a temporary abode of purchasing power. Thus concludes Friedman, "The selection of a specific empirical counterpart to the term money seems to us a matter of convenience for a particular purpose, not a matter of principle."[6] He is, therefore, not rigid in his definition of money and takes a broader view which includes *bank deposits, non-bank deposits and any other type of assets through which the monetary authority influences the future level of income, prices, employment or any other important macro variable.*

3. The Radcliffe Definition. The Radcliffe Committee[7] defined money as "notes plus bank deposits". It includes as money only those assets which are commonly used as media of exchange. Assets refer to liquid assets by which it means the monetary quantity influencing total effective demand for goods and services. This is interpreted widely to include credit. Thus the whole liquidity position is relevant to spending decisions. Spending is not limited to cash or money in the bank but to the amount of money people think they can get hold of either by selling an asset or by borrowing or by receipts of income from, say, sales. The Committee did not make use of the concept of velocity of circulation because as a numerical constant, it is devoid of any behavioural content. On the basis of crude empirical tests, the Committee did not find either direct or indirect link between money and economic activity via the interest rate. But it gave a new transmission mechanism based on liquidity. It explained that a movement of interest rates implies significant changes in the capital value of many assets held by financial institutions. A rise in the interest rates makes some less willing to lend because capital values have fallen, and others because their own interest rate structure is sticky. A fall in interest rates, on the other hand, strengthens balance sheets and encourages lenders to seek new business.

4. The Gurley-Shaw Definition. Gurley and Shaw[8] regard a substantial volume of liquid assets held by financial intermediaries and the liabilities of non-bank intermediaries as close substitutes for money. Intermediaries provide substitutes for money as a store of value. Money proper which is defined as equal to currency plus demand deposits is only one liquid asset. They have thus formulated a wider definition of money based upon liquidity which includes bonds, insurance reserves, pension funds, savings and loan shares. They believe in the velocity of the money stock which is influenced by non-bank intermediaries. Their views on the definition of money are based on their own and Goldsmith's empirical findings.

---

6. M. Friedman and A.J. Schwartz, *Monetary Statistics of the United States*, 1970.

7. Report of the Committee on the Working of Monetary System, 1959.

8. J.G. Gurley and E.S. Shaw, *Money in a Theory of Finance*, 1960.

5. The Pesek and Saving Definition. According to Pesek and Saving,[9] money should include demand deposits of banks as well as money issued by the government. They exclude time and saving deposits from bank money. They regard total money which includes demand deposits as net wealth of society. They contrast money with debt. Money does not pay interest but debt yields interest. Debt itself is not wealth because those who hold bank money consider it as an asset while banks consider it as an effective liability. Thus Pesek-Saving follow a usable definition of money which consists of three conditions: *First,* they regard commodity money and fiat money as assets to their holders and liabilities to no one. *Second,* the government grants monopoly rights to commercial banks to produce money who, in turn, exercise it by selling bank money for the private debts of individuals. Individuals who hold bank money fully consider it as an asset. On the other hand, banks consider it as an effective liability. Thus Pesek-Saving regard bank money less reserves (which banks hold to meet the demand of their depositors) as a net asset of the economy. In the balance sheet, bank money is shown as an asset and private debts as a liability. *Third,* if it is costless to produce bank money and no interest payments are made on deposits, the net worth of the bank remains unchanged because both assets and liabilities increase by the same amount. This shows that the bank has zero net worth.

Pesek-Saving do not include time and saving deposits in bank money. But when interest payments are involved, they are included in bank money. They argue that once these deposits start paying interest, they will continue to serve as money.

According to Friedman and Schwartz, the logic of Pesek-Saving position is that money as net wealth ought to be high-powered money which the latter rejected as too narrow. It is thus high-powered money or the monetary base which satisfies Pesek-Saving first criterion.

Patinkin finds some confusion in Pesek-Saving analysis when they exclude time and saving deposits from bank money. Pesek-Saving also criticise the traditional methods of accounting used by banks. But Patinkin explains their analysis of money creation in traditional accounting and finds it more useful.

Pesek-Saving have been criticised for double counting bank money in defining social wealth. First, they include it as part of the money supply and then they include the net worth of the banking system in another entity. In fact, one or the other should be counted, and not both.

Despite these criticisms, the views of Pesek and Saving on money are important because they study net worth which accrues to commercial banks.

MONEY AND NEAR MONEY

Money consists of currency and bank deposits. Coins and currency notes issued by the central bank of a country and cheques of commercial banks of a country are liquid assets. In fact, cheques and bank drafts are almost perfect substitutes for money. This is because they perform the medium of exchange function of money. But cheques and drafts can be issued at a short notice only in the case of *demand deposits*. This is not the case with time deposits which can be withdrawn either at the end of the fixed period or by giving a prior notice to the bank and incurring a penalty. Thus time deposits are not 'real' money and for them to become money they must be converted into cash or demand deposits. However, they are near money for they can be converted into real money in a short period without any loss. For example, a time deposit of Rs. 10,000 for three years can always be converted into demand deposit or Rs. 10,000 in cash at any time before the completion of three years. They also fetch a fixed rate of interest which may be reduced if the deposit is converted into cash or demand deposit before the expiry of the fixed period. Thus near money assets serve the store of value function of money temporarily and are convertible into a medium of exchange in a short time without loss in their face value.

9. B.P. Pesek and T.R. Saving, *Money, Wealth and Economic Theory*, 1967.

Besides time deposits, other near money assets are bonds, securities, debentures, bills of exchange, treasury bills, insurance policies, etc. All these types of assets have a market and are negotiable so that they can be converted into real money within a short time. We discuss how these negotiable instruments are near money.

Bonds, securities and debentures fall in the same category. Bonds and securities are issued by the government, while debentures are issued by companies. They are the means to borrow funds for short, medium or long periods and carry a fixed rate of interest. They are near money assets because they are convertible into cash at a short notice in the money market.

A bill of exchange is another form of money. It is an IOU (I owe you). It is drawn by an individual or firm to pay a stated sum of money on a specified date which is never more than 90 days. A bill of exchange is not money by itself but is certainly money on the due date. It is, however, near money if its owner wishes to turn it into cash. It can be easily converted into money at a discount or by receiving less money than its face value.

Treasury bills issued by the government also fall in the category of near money. A treasury bill is a promise by the government to pay a stated sum in the near future usually 30, 60 or 90 days. A treasury bill is also like a bill of exchange which is convertible into money at a discount within a short period.

Life insurance policy is another example of near money. The holder of a life insurance policy can obtain cash in the form of loan on his policy at a short notice. Thus a life insurance policy is a form of liquid asset which can be regarded as near money.

Besides these recognised instruments of near money, some intermediaries have come into existence to provide a market for certain assets. Such intermediaries are financial companies which provide funds on the security of some assets, and brokers who buy and sell property, bonds, debentures, shares, etc. They help in increasing the liquidity of such assets thereby converting them into near money.

Money and near money can now be distinguished. Money is a legal-tender and gives the possessor liquidity in hand. It performs the medium of exchange function. On the other hand, near money assets do not have any legal status. They possess moneyness or liquidity but not ready liquidity like money. They are almost perfect substitutes for money as a store of value. They are superior to money because they yield income. They also

economise in the use of money proper and tend to reduce the quantity of money used by the people as a medium of exchange, as a medium of deferred payments and as a store of value.

Despite the fact that near money assets do not possess ready liquidity, they are preferred by individuals. According to Professor A.G. Hart, near money is preferred to cash by individuals because it serves as a margin of safety motive. Prof. Dean points out that 80 per cent of near money in the USA is held by individuals. The tremendous growth of near money assets in the United States is due to the fact that the yield from them is higher than from demand deposits and that they are safer than cash. Moreover, savings in the near money assets have been encouraged by such special techniques as pay roll saving plan, and commercial bank plans covering automatic regular transfer to specified sums from the demand to time deposits of the customers.

INSIDE MONEY AND OUTSIDE MONEY

Gurley and Shaw[10] have divided the total quantity of money in the economy into two parts : inside money and outside money. They define *inside money* as money "based on the debt of endogenous economic units". It is money created within the private sector against some private debt. It consists of financial claims by one class of individuals and firms on others within the economy. When banks create deposits through loans and advances, these return to them as secondary deposits. Deposits and loans represent assets and liabilities of the private sector. On the other hand, *outside money* comes from outside the private sector. It is defined as money "based on the debt of a unit (the government) exogenous to the system." Thus it is currency issued by the government, gold, foreign and government securities. It is the liability of the government as a debtor and claim of the private sector as a creditor.

The distinction between inside money and outside money is important from many viewpoints.

First, we study the importance of the distinction between the two in a period of rising or falling prices. When there is inflation or deflation, a

change in the purchasing power of money in the case of *inside money* leads to an equal change in the real values of both assets and liabilities of the private sector. Thus a change in the price level does not affect the behavior of depositors and borrowers possessing inside money. The monetary system also does not gain or lose in real terms by such changes in the real amount of its debt because there is an equal change in the real value of its claims against firms. Given the nominal amount of inside money, its real value varies inversely with the price level. In other words, a change in the price level does not lead to a wealth transfer between the private sector and the government. Instead it results only in a wealth transfer between consumers and firms, the former gaining and the latter losing when the price level falls and vice versa. This transfer is a distribution effect of instability in the price level. But such a change in the price level, when money is of the inside variety, does not affect government behaviour and has no net effect on total wealth in the private sector.

G-S have shown that in this system, the price level is determinate. A departure from the initial equilibrium price level sets in motion forces which tend to restore the initial level or a new equilibrium price level. For instance, when the price level rises, it disturbs the portfolio equilibrium of the private sector by decreasing its real money holdings relative to its bond holdings. At the new price level, there will be excess real demand for money, excess real supply of bonds and goods, so that the system reverts back to the initial price level and the initial stock of nominal bonds. This will induce sales in the bond market to raise the interest rate and expenditure in the commodity market to raise the price level. The increase in the interest rate would have a depressing effect through commodity market, thereby bringing back the initial price level, and vice versa in case of deflation.

---

10. J.G. Gurley and E.S. Shaw, *Money in a Theory of Finance*, 1960.

In the case of *outside money* when prices rise, the real value of cash held by exogenous economic units falls. On the contrary, when prices fall, the real value of cash held by economic units rises. Thus a change in the price level affects the behaviour of economic units possessing outside money.

Thus given the nominal amount of outside money, its real value varies inversely with the price level and each change in its real value leads to a wealth transfer between the private sector and the government. This wealth transfer affects private demands for money, goods and labour but not the government demand.

When money is of outside variety, its impact is neutral on the real variables of the economy. A monetary change causes an equi-proportionate change in prices. As a result, relative prices and interest rate would remain unchanged. An expansion in the nominal stock of outside money causes excess supply in a money market and excess demand in the commodity market. This leads to the rise in price level above its initial level. The price rise reduces the aggregate real wealth of the private sector and a decline in the real value of government debt. This brings down the real demands in all markets which restore the initial price level. In this process, there is a wealth transfer between the private sector and the government which has a *net* effect on aggregate demands for money, goods and labour

NEUTRALITY AND NON-NEUTRALITY OF MONEY

Neutrality of Money. Neutrality of money means that money is neutral in its effect on the economy. A change in the money stock can have no long-run influences on the level of real output, employment, rate of interest, or the composition of final output. The only lasting impact of a change in the money stock is to alter the general price level. Patinkin explains the neutrality of money as a situation when “a uniformly introduced increase in the quantity of money causes a proportionate increase in the equilibrium price of commodities and leaves the equilibrium rate of interest unaffected,” provided there is absence of money illusion and distribution effects.[11] According to Gurley and Shaw, money is neutral if money is either entirely of the "outside" variety, or entirely of the "inside variety. They define neutrality of money as the "inability of changes in the nominal stock of money to affect the rate of interest, output and wealth, and other variables."

In other words, money is neutral if it does not affect relative prices and

leaves the interest rate unaffected. All prices move equiproportionally. If this happens without a time lag, the neutrality of money is instantaneous. If there is a time lag, there is long-run neutrality. These are changes in absolute (money) prices but individual economic units are unresponsive to them. The quantity of money determines only absolute prices and their level does not affect the level of income, interest, rate of capital formation and employment. It is in this sense that money is neutral in its effects on the working of the economy.

In the *classical system*, money is neutral in its effect on the economy. It plays no role in the determination of employment, income and output. Rather, they are determined by labour, capital stock, state of technology, availability of natural resources, saving habits of the people, and so on. In the classical system, the main function of money is to act as a medium of exchange. It is to determine the general level of prices at which goods and services will be exchanged. The quantity theory of money states that price level is a function of the supply of money. Algebraically, $MV = PT$, where M, V, P and T are the supply of money, velocity of money, price level and the volume of transactions (or total output) respectively. The equation tells that the total money supply, MV, equals the total value of output, PT, in the economy. Assuming V and T to be constant, a change in M causes a proportionate change in P. Thus money is neutral whose main function is to determine the general price level at which goods and services exchange.

---

11. D. Patinkin, *Interest and Prices*, 2/e, 1965.

In the entire *Keynesian system*, there are two situations in which money is neutral. The *first* is the situation of full employment when any increase in the quantity of money brings about a proportionate increase in the price level but output remains unchanged at that level. The *second* is the special case of liquidity trap. When the economy is in the liquidity trap, there cannot be a further fall in the rate of interest even if the money supply is increased by monetary authorities. This implies that there will be no effect on such real variables as investment and income and all changes in the money supply are added to idle balances. In this situation, money is neutral. However, money is non-neutral in the intermediate situation between these two extreme cases in the Keynesian system.

Non-Neutrality of Money. In the Keynesian system so long as there is unemployment, changes in the money supply produce permanent non-neutral effects on the rate of interest, the level of employment, income and output, the rate of capital formation, and so on. Thus Keynes emphasized non-neutral money and for this he invoked the monetary theory of interest. He wrote: "As soon as we pass to the problem of what determines output and employment as a whole, we require the complete theory of a Monetary Economy."

The post-Keynesians, particularly Friedman, Burner and Metzler have shown that money is non-neutral in the short-run. According to Friedman, interest rates cannot be used as a guide to monetary policy and that an acceleration in the rate of growth of the money supply produces not lower interest rates but higher ones, if the entire cycle of events is considered. Thus monetary policy cannot peg interest rates except for limited periods. Any attempt to do so will only produce sustained inflation, just as would a similar attempt to use monetary policy to hold unemployment below its natural rate."

Besides, Friedman also believes that money may be non-neutral in the long-run. If there is a permanent acceleration in the growth rate of the money supply, say from 3 per cent to 8 per cent, it will permanently change the level of real income. This implies non-neutrality of money. This view presupposes that (*i*) real money balances are a productive factor to business and provider of utility to the wealthholder, and (*ii*) inflation reduces the real balances with business and wealth-holder.

FUNCTIONS OF MONEY

Money performs a number of primary, secondary, contingent and other functions which not only remove the difficulties of barter but also oils the wheels of trade and industry in the present day world. We discuss these functions one by one.

1. Primary Functions

The two primary functions of money are to act as a medium of exchange

and as a unit of value.

(i) Money as a Medium of Exchange. This is the primary function of money because it is out of this function that its other functions developed. By serving as a medium of exchange, money removes the need for double coincidence of wants and the inconveniences and difficulties associated with barter. The introduction of money as a medium of exchange decomposes the single transaction of barter into separate transactions of sale and purchase, thereby eliminating the double coincidence of wants. This function of money also separates the transactions in time and place because the sellers and buyers of a commodity are not required to perform the transactions at the same time and place. This is because the seller of a commodity buys some money and money, in turn, buys the commodity over time and place.

When money acts as a medium of exchange, it means that it is generally acceptable. It, therefore, affords the freedom of choice. With money, we can buy an assorted bundle of goods and services. At the same time, we can purchase the best and also bargain in the market. Thus money gives us a good deal of economic independence and also perfects the market mechanism by increasing competition and widening the market.

As a medium of exchange, money acts as an intermediary. It facilitates exchange. It helps production indirectly through specialisation and division of labour which, in turn, increase efficiency and output. According to Prof. Walters, money serves as a 'factor of production,' enabling output to increase and diversify.

In the last analysis money facilitates trade. When acting as the intermediary, it helps one good or service to be traded indirectly for others.

(ii) Money as Unit of Value.[12] The second primary function of money is to act as a unit of value. Under barter one would have to resort to some standard of measurement, such as a length of string or a piece of wood. Since one would have to use a standard to measure the length or height of any object, it is only sensible that one particular standard should be accepted as *the* standard. Money is the standard for measuring value just

as the yard or metre is the standard for measuring length. The monetary unit measures and expresses the values of all goods and services. In fact, it expresses the value of each good or service in terms of price. Money is the common denominator which determines the rate of exchange between goods and services which are priced in terms of the monetary unit. There can be no pricing process without a measure of value.

The use of money as a standard of value eliminates the necessity of quoting the price of apples in terms of oranges, the price of oranges in terms of nuts and so on. Unlike barter, the prices of such commodities are expressed in terms of so many units of dollars, rupees, francs, pounds, etc., depending on the nature of the monetary unit in a country. As a matter of fact, measuring the values of goods and services in the monetary unit facilitates the problem of measuring the exchange values of goods in the market. When values are expressed in terms of money, the number of prices are reduced from n(n–1) in barter economy to (n–1) in monetary economy.

---

12. This function is known by various other names such as unit of account, standard of value, common measure of value or common denominator of value.

Money as a unit of value also facilitates accounting when assets of all kinds, liabilities of all kinds, incomes of all kinds, and expenses of all kinds can be stated in terms of common monetary units to be added or substracted.

Further, money as a unit of account helps in calculations of economic importance such as the estimation of the costs, and revenues of business firms, the relative costs and profitability of various public enterprises and projects under a planned economy, and the gross national product. As pointed out by Culbertson, "Prices quoted in terms of money become the focus of people's behaviour. Their calculations, plans, expectations, and contracts focus on money prices."[13]

2. Secondary Functions

Money performs three secondary functions: as a standard of deferred

payments, as a store of value, and as a transfer of value. They are discussed below.

(i) Money as a Standard of Deferred Payments. The third function of money is that it acts as a standard of deferred or postponed payments. All debts are taken in money. It was easy under barter to take loans in goats or grains but difficult to make repayments in such perishable articles in the future. Money has simplified both the taking and repayment of loans because the unit of account is durable. Money links the present values with those of the future. It simplifies credit transactions. It makes possible contracts for the supply of goods in the future for an agreed payment of money. It simplifies borrowing by consumers on hire-purchase and from house-building and cooperative societies. Money facilitates borrowing by firms and businessmen from banks and non-bank financial institutions. The buying and selling of shares, debentures and securities is made possible by money. By acting as a standard of deferred payments, money helps in capital formation both by the government and business enterprises. In fine, this function of money develops financial and capital markets and helps in the growth of the economy.

But there is the danger of changes in the value of money over time which harm or benefit the creditors and debtors. If the value of money increases over time, the creditors gain and debtors lose. On the other hand, a fall in the value of money over time brings losses to creditors and windfalls to debtors.

(ii) Money as a Store of Value. Another important function of money is that it acts as a store of value. "The good chosen as money is always something which can be kept for long periods without deterioration or wastage. It is a form in which wealth can be kept intact from one year to the next. Money is a bridge from the present to the future. It is therefore essential that the money commodity should always be one which can be easily and safely stored."[14]

Money as a store of value is meant to meet unforeseen emergencies and to pay debts. Newlyn calls this the *asset function of money*. "Money is not, of course, the only store of value. This function can be served by any valuable asset. One can store value for the future by holding short-term

promissory notes, bonds, mortgages, preferred stocks, household furniture, houses, land, or any other kind of valuable goods. The principal advantages of these other assets as a store of value are that they, unlike money, ordinarily yield an income in the form of interest, profits, rent or usefulness...,and they sometimes rise in value in terms of money. On the other hand, they have certain disadvantages as a store of value, among which are the following: (1) They sometimes involve storage costs; (2) they may depreciate in terms of money; and (3) they are "illiquid" in varying degrees, for they are not generally acceptable as money and it may be possible to convert them into money quickly only by suffering a loss of value."[15]

---

13. J.M. Culbertson, Money and Banking, 1972.

14. A.W. Stonier and D.C. Hague, *op.cit.*, p. 406.

Keynes placed much emphasis on this function of money. According to him, to hold money is to keep it as a reserve of liquid assets which can be converted into real goods. It is a matter of comparative indifference whether wealth is in money, money claims, or goods. In fact, money and money claims have certain advantages of security, convenience and adaptability over real goods. But the store of value function of money also suffers from changes in the value of money. This introduces considerable hazard in using money or assets as a store of value.

(iii) Money as a Transfer of Value. Since money is a generally acceptable means of payment and acts as a store of value, it keeps on transferring values from person to person and place to place. A person who holds money in cash or assets can transfer that to any other person. Moreover, he can sell his assets at Delhi and purchase fresh assets at Bangalore. Thus money facilitates transfer of value between persons and places.

3. Contingent Functions

Money also performs certain contingent or incidental functions, according to Prof. David Kinley. They are :

(i) Money as the Most Liquid of all Liquid Assets. Money is the most liquid of all liquid assets in which wealth is held. Individuals and firms may hold wealth in infinitely varied forms. They may, for example, choose between holding wealth in currency, demand deposits, time deposits, savings, bonds, treasury bills, short-term government securities, long-term government securities, debentures, preference shares, ordinary shares, stocks of consumer goods, and productive equipment. All these are liquid forms of wealth which can be converted into money, and vice-versa.

(ii) Basis of the Credit System. Money is the basis of the credit system. Business transactions are either in cash or on credit. Credit economises the use of money. But money is at the back of all credit. A commercial bank cannot create credit without having sufficient money in reserve. The credit instruments drawn by businessmen have always a cash guarantee supported by their bankers.

(iii) Equaliser of Marginal Utilities and Productivities. Money acts as an equaliser of marginal utilities for the consumer. The main aim of a consumer is to maximise his satisfaction by spending a given sum of money on various goods which he wants to purchase. Since prices of goods indicate their marginal utilities and are expressed in money, money helps in equalising the marginal utilities of various goods. This happens when the ratios of the marginal utilities and prices of the various goods are equal.

(iv) Equaliser of Marginal Productivities. Money also helps in equalising the marginal productivities of the various factors. The main aim of the producer is to maximise his profits. For this, he equalises the marginal productivity of each factor with its price. The price of each factor is nothing but the money he receives for his work.

(v) Measurement of National Income. It was not possible to measure national income under the barter system. Money helps in measuring national income. This is done when various goods and services produced in a country are assessed in money terms.

(vi) Distribution of National Income. Money also helps in the

distribution of national income. Rewards of factors of production in the form of wages, rent, interest and profit are determined and paid in terms of money.

4. Other Functions

Money also performs such functions which affect the decisions of consumers and governments.

(i) Helpful in Making Decisions. Money as a means of store of value helps the consumer meet his daily requirements on the basis of money held by him. If the consumer has a scooter and he needs a car in the near future, he can buy a car by selling his scooter and money accumulated by him. In this way, money helps in taking decisions.

(ii) Money as a Basis of Adjustment. To carry on trade in a proper manner, the adjustment between money market and capital market is done through money. Similarly, adjustments in foreign exchange are also made through money. Further, international payments of various types are also adjusted and made through money. It is on the basis of these functions that money guarantees the solvency of the payer and provides options to the holder of money to use it any way, he likes.

---

15. *W.T. Newlyn,* Theory of Money, *2/e, Ch. 1*

SELECTED READINGS

L.V. Chandler and S.M. Goldfield. *The Economics of Money and Banking,* 7/e, 1977.

G. Growther, *An Outline of Money,* 1958.

W.T. Newlyn, *Theory of Money,* 2/e, Ch. 1.

G.J. Gurley and S.E. Shaw, *Money in a Theory of Finance,* 1960.

EXERCISES

1. Discuss the meaning and nature of money.
2. “Money is what money does”. How far do you agree with this definition of money?
3. How is money different from other liquid assets?
4. Distinguish between money and near money assets.
5. Distinguish between inside money and outside money.
6. Explain neutrality of money. When is money non-neutral?
7. Briefly explain the functions of money.

CHAPTER 25

CHANGES IN THE VALUE OF MONEY: THE QUANTITY THEORY OF MONEY AND ITS VARIANTS

MEANING OF VALUE OF MONEY

By value of money is meant the purchasing power of money over goods and services in a country. What a rupee can buy in India represents the value of money of the rupee. Thus the phrase, "value of money" is a relative concept which expresses the relationship between a unit of money and the goods and services which can be purchased with it. This shows that the value of money is related to the price level because goods and services are purchased with a money unit at given prices. But the relation between the value of money and price level is an inverse one. If V presents the value of money and P the price level, then, $V = 1/P$. When the price level rises, the value of money falls, and vice versa. Thus in order to measure the value of money, we have to find out the general price level.

The value of money is of two types: The internal value of money and the external value of money. The internal value of money refers to the purchasing power of money over domestic goods and services. The external value of money refers to the purchasing power of money over foreign goods and services.

FISHER'S QUANTITY THEORY OF MONEY: THE CASH TRANSACTIONS APPROACH

The quantity theory of money states that the quantity of money is the main determinant of the price level or the value of money. Any change in the quantity of money produces an exactly proportionate change in the price level. In the words of Irving Fisher, "Other things remaining unchanged, as the quantity of money in circulation increases, the price level also increases in direct proportion and the value of money decreases and vice versa." If the quantity of money is doubled, the price level will also double and the value of money will be one half. On the other hand, if the quantity of money is reduced by one half, the price level will also be reduced by one half and the value of money will be twice.

Fisher has explained his theory in terms of his equation of exchange:

$$PT = MV + M'V'$$

where P = price level, or $1/P$ = the value of money;

M = the total quantity of legal tender money;

V = the velocity of circulation of M;

M' = the total quantity of credit money;

V' = the velocity of circulation of M';

T = the total amount of goods and services exchanged for money or transactions performed by money.

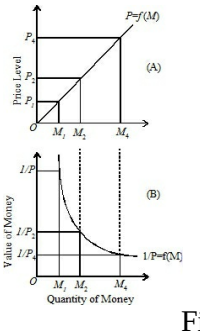


Fig 1.

This equation equates the demand for money (PT) to supply of money ($MV=M'V'$). The total volume of transactions multiplied by the price level (PT) represent the demand for money. According to Fisher, PT is ΣPQ. In other words, price level (P) multiplied by quantity bought (Q) by the community (Σ) gives the total demand for money. This equals the total supply of money in the community consisting of the quantity of actual money M and its velocity of circulation V *plus* the total quantity of

credit money M' and its velocity of circulation V'. Thus the total value of purchases (PT) in a year is measured by $MV+M'V'$. Thus the equation of exchange is $PT=MV+M'V'$. In order to find out the effect of the quantity of money on the price level or the value of money, we write the equation as

$$P = \frac{MV+M'V'}{T}$$

Fisher points out that the price level (P) varies directly as the quantity of money ($M+M'$), provided the volume of trade (T) and velocity of circulation (V, V') remain unchanged. The truth of this proposition is evident from the fact that if M and M' are doubled, while V, V' and T remain constant, P is also doubled, but the value of money ($1/P$) is reduced to half.

Fisher's quantity theory of money is explained with the help of Figure 1 (A) and (B). Panel A of the figure shows the effect of changes in the quantity of money on the price level. To begin with, when the quantity of money is M_1, the price level is P_1. When the quantity of money is doubled to M_2, the price level is also doubled to P_2. Further, when the quantity of money is increased four-fold to M_4, the price level also increases by four times to P_4. This realtionship is expressed by the curve $P=f(M)$ from the origin at 45° .

In Panel B of the figure, the inverse relation between the quantity of money and the value of money is depicted where the value of money is taken on the vertical axis. When the quantity of money is M_1, the value of money is $1/P$. But with the doubling of the quantity of money to M_2, the value of money becomes one-half of what it was before, $1/P_2$. And with the quantity of money increasing by four-fold to M_4, the value of money is reduced by $1/P_4$. This inverse relationship between the quantity of money and the value of money is shown by downward sloping curve $1/P=f(M)$.

Assumptions of the Theory

Fisher's theory is based on the following assumptions:

1. *P* is a passive factor in the equation of exchange which is affected by the other factors.

2. The proportion of *M'* to *M* remains constant.

3. *V* and *V'* are assumed to be constant and are independent of changes in *M* and *M'*.

4. *T* also remains constant and is independent of other factors such as *M*, *M'*, *V* and *V'*.

5. It is assumed that the demand for money is proportional to the value of transactions.

6. The supply of money is assumed as an exogenously determined constant.

7. The theory is applicable in the long run.

8. It is based on the assumption of the existence of full employment in the economy.

Criticisms of the Theory

The Fisherian quantity theory has been subjected to severe criticisms by economists.

1. Truism. According to Keynes, "The quantity theory of money is a truism." Fisher's equation of exchange is a simple truism because it states that the total quantity of money ($MV+M'V'$) paid for goods and services must equal their value (PT). But it cannot be accepted today that a certain percentage change in the quantity of money leads to the same percentage change in the price level.

2. Other Things not Equal. The direct and proportionate relation between quantity of money and price level in Fisher's equation is based on the assumption that "other things remain unchanged". But in real life, *V*, *V'* and *T* are not constant. Moreover, they are not independent of *M*, *M'*

and *P*. Rather, all elements in Fisher's equation are interrelated and interdependent. For instance, a change in *M* may cause a change in *V*.

Consequently, the price level may change more in proportion to a change in the quantity of money. Similarly, a change in *P* may cause a change in *M*. Rise in the price level may necessitate the issue of more money. Moreover, the volume of transactions *T* is also affected by changes in *P*. When prices rise or fall, the volume of business transactions also rises or falls. Further, the assumptions that the proportion *M'* to *M* is constant, has not been borne out by facts. Not only this, *M* and *M'* are not independent of *T*. An increase in the volume of business transactions requires an increase in the supply of money (*M* and *M'*).

3. Constants Relate to Different Time. Prof. Halm criticises Fisher for multiplying *M* and *V* because *M* relates to a point ot time and *V* to a period of time . The former is a static concept and the latter a dynamic. It is, therefore, technically inconsistent to multiply two non-comparable factors.

4. Fails to Measure Value of Money. Fisher's equation does not measure the purchasing power of money but only cash transactions, that is, the volume of business transactions of all kinds or what Fisher calls the volume of trade in the community during a year. But the purchasing power of money (or value of money) relates to transactions for the purchase of goods and services for consumption. Thus the quantity theory fails to measure the vlaue of money.

5. Weak Theory. According to Crowther, the quantity theory is weak in many respects. *First*, it cannot explain 'why' there are fluctutations in the price level in the short run. *Second*, it gives undue importance to the price level as if changes in prices were the most critical and important phenomenon of the economic system. *Third*, it places a misleading emphasis on the quantity of money as the principal cause of changes in the price level during the trade cycle. Prices may not rise despite increase in the quantity of money during depression; and they may not decline with reduction in the quantity of money during boom. Further, low prices during depression are not caused by shortage of quantity of money, and high prices during prosperity are not caused by abundance of quantity of money. Thus, "the quantity theory is at best an imperfect guide to the

causes of the trade cycle *in the short period*," according to Crowther.

6. Neglects Interest Rate. One of the main weaknesses of Fisher's quantity theory of money is that it neglects the role of the rate of interest as one of the causative factors between money and prices. Fisher's equation of exchange is related to an equilibrium situation in which rate of interest is independent of the quantity of money.

7. Unrealistic Assumptions. Keynes in his *General Theory* severely criticised the Fisherian quantity theory of money for its unrealistic assumptions. *First*, the quantity theory of money is unrealistic because it analyses the relation between *M* and *P* in the long run. Thus it neglects the short run factors which influence this relationship. *Second*, Fisher's equation holds good under the assumption of full employment. But Keynes regards full employment as a special situation. The general situation is one of the underemployment equilibrium. *Third*, Keynes does not believe that the relationship between the quantity of money and the price level is direct and proportional. Rather, it is an indirect one via the rate of interest and the level of output. According to Keynes, "So long as there is unemployment, *output* and employment will change in the same proportion as the quantity of money, and when there is full employment, prices will change in the same proportion as the quantity of money." Thus Keynes integrated the theory of output with value theory and monetary theory and criticised Fisher for dividing economics "into two compartments with no doors and windows between *the theory of value and theory of money and prices*."

8. V not Constant. Further, Keynes pointed out that when there is underemployment equilibrium, the velocity of circulation of money *V* is highly unstable and would change with changes in the stock of money or money income. Thus it was unrealistic for Fisher to assume *V* to be constant and independent of *M*.

9. Neglects Store of Value Function. Another weakness of the quantity theory of money is that it concentrates on the supply of money and assumes the demand for money to be constant. In order words, it neglects the store-of-value function of moeny and considers only the medium-of-exchange function of money. Thus the theory is one-sided.

10. Neglects Real Balance Effect. Don Patinkin has critcised Fisher for failure to make use of the real balance effect, that is, the real value of cash balances. A fall in the price level raises the real value of cash balances which leads to increased spending and hence to rise in income, output and employment in the economy. According to Patinkin, Fisher gives undue importance to the quantity of money and neglects the role of real money balances.

11. Static. Fisher's theory is static in nature because of its such unrealistic assumptions as long run, full employment, etc. It is, therefore, not applicable to a modern dynamic economy.

THE CAMBRIDGE EQUATIONS : THE CASH BALANCES APPROACH

As an alternative to Fisher's quantity theory of money, Cambridge economists Marshall, Pigou, Robertson and Keynes formulated the cash balances approach. Like value theory, they regarded the determination of value of money in terms of supply and demand. Robertson wrote in this connection: "Money is only one of the many economic things. Its value, therefore, is primarily determined by exactly the same two factors as determine the value of any other thing, namely, the conditions of demand for it, and the quantity of it available."

The supply of money is exogenously determined at a point of time by the banking system. Therefore, the concept of velocity of circulation is altogether discarded in the cash balances approach because it 'obscures the motives and decisions of people behind it'. On the other hand, the concept of demand for money plays the major role in determining the value of money. The demand for money is the demand to hold cash balances for transactions and precautionary motives.

Thus the cash balances approach considers the demand for money not as a medium of exchange but as a store of value. Robertson expressed this distinction as money "on the wings" and money "sitting". It is "money sitting" that reflects the demand for money in the Cambridge equations. The Cambridge equations show that given the supply of money at a point of time, the value of money is determined by the demand for cash

balances. When the demand for money increases, people will reduce their expenditures on goods and services in order to have larger cash holdings. Reduced demand for goods and services will bring down the price level and raise the value of money. On the contrary, fall in the demand for money will raise the price level and lower the value of money.

The Cambridge cash balances equations of Marshall, Pigou, Robertson and Keynes are discussed as under:

Marshall's Equation. Marshall did not put his theory in equation form and it was for his followers to explain it algebraically. Friedman has explained Marshall's views thus: "As a first approximation, we may suppose that the amount one wants to hold bears some relation to one's income, since that determines the volume of purchases and sales in which one is engaged. We then add up the cash balances held by all holders of money in the community and express the total as a fraction of their total income." Thus we can write:

$$M = kPY$$

where M stands for the exogenously determined supply of money, k is the fraction of the real money income (PY) which people wish to hold in cash and demand deposits, P is the price level, and Y is the aggreagate real income of the community. Thus the price level $P = \frac{M}{kY}$ or the value of money (the reciprocal of price level) is $\frac{1}{P} = \frac{kY}{M}$

Pigou's Equation. Pigou was the first Cambridge economist to express the cash balances approach in the form of an equation:

$$P = \frac{kR}{M}$$

where P is the purchasing power of money or the value of money (the

reciprocal of the price level), *k* is the proportion of total real resources or income (*R*) which people wish to hold in the form of titles to legal tender, *R* is the total resources (expressed in terms of wheat), or real income, and *M* refers to the number of actual units of legal tender money.

The demand for money, according to Pigou, consists not only of legal money or cash but also bank notes and bank balances. In order to include bank notes and bank balances in the demand for money, Pigou modifies his equation as

$$P = \frac{kR}{M}\{c + h(1-c)\}$$

where *c* is the proportion of total real income actually held by people in legal tender including token coins, (1-*c*) is the proportion kept in bank notes and bank balances, and *h* is the proportion of actual legal tender that bankers keep against the notes and balances held by their customers.

Pigou points out that when *k* and *R* in the equation *P=kR/M* and *k*, *R*, *c* and *h* are taken as constants then the two equations give the demand curve for legal tender as a rectangular hyperbola. This implies that the demand curve for money has a uniform unitary elasticity. This is shown in Figure 2 where DD_1 is the demand curve for money and Q_1M_1, Q_2M_2, and Q_3M_3 are the supply curves of money drawn on the assumption that the supply of money is fixed at a point of time. The value of money or Pigou's purchasing power of money *P* is taken on the vertical axis. The figure shows that when the supply of money increases from OM_1 to OM_2, the value of money is reduced from OP_1 to OP_2. The fall in the value of money by P_1P_2 exactly equals the increase in the supply of money by M_1M_2. If the supply of money increases three times from OM_1 to OM_3, the value of

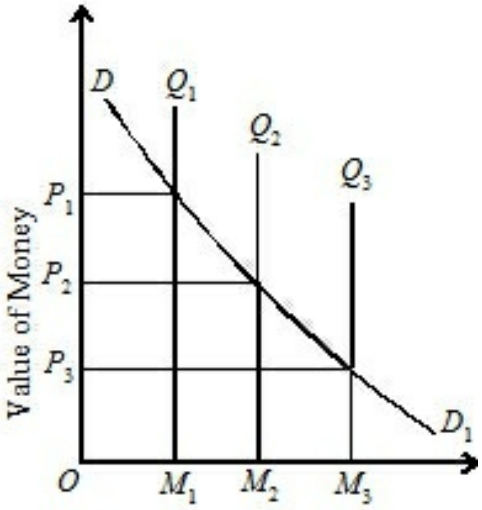


Fig. 2.

money is reduced by exactly one-third from OP_1 to OP_3. Thus the demand curve for money DD_1 is a rectangular hyperbola because it shows changes in the value of money exactly in reverse proportion to the supply of money.

Robertson's Equation. To determine the value of money or its reciprocal the price level, Robertson formulated an equation similar to that of Pigou. The only difference between the two being that instead of Pigou's total real resources *R*, Robertson gave the volume of total transactions *T*. The Robertsonian equation is $M = PkT$ or

$$P = \frac{M}{kT}$$

where *P* is the price level, *M* is the total quantity of money, *k* is the proportion of the total amount of goods and services (*T*) which people wish to hold in the form of cash balances, and *T* is the total volume of goods and services purchased during a year by the community.

If we take *P* as the value of money instead of the price level as in Pigou's equation, then Robertson's equation exactly resembles Pigou's $P = kT/M$.

Keynes's Equation. Keynes in his *A Tract on Monetary Reform* (1923) gave his Real Balances Quantity Equation as an improvement over the other Cambridge equations. According to him, people always want to have some purchasing power to finance their day to day transactions. The amount of purchasing power (or demand for money) depends partly on their tastes and habits, and partly on their wealth. Given the tastes, habits, and wealth of the people, their desire to hold money is given. This demand for money is measured by consumption units. A consumption unit is expressed as a basket of standard articles of consumption or other objects of expenditure.

If *k* is the number of consumption units in the form of cash, *n* is the total currency in circulation, and *p* is the price for consumption unit, then the equation is

$$n = pk$$

If k is constant, a proportionate increase in n (quantity of money) will lead to a proportionate increase in p (price level).

This equation can be expanded by taking into account bank deposits. Let k' be the number of consumption units in the form of bank deposits, and r the cash reserve ratio of banks, then the expanded equation is

$$n = p\,(k + rk')$$

Again, if k, k' and r are constant, p will change in exact proportion to the change in n.

Keynes regards his equation superior to other cash balances equations. The other equations fail to point how the price level (p) can be regulated. Since the cash balances (k) held by the people are outside the control of the monetary authority, p can be regulated by controlling n and r. It is also possible to regulate bank deposits k' by appropriate changes in the bank rate. So p can be controlled by making appropriate changes in n, r and k' so as to offset changes in k.

Criticisms of the Cash Balance Approach

The cash balances approach to the quantity theory of money has been criticised on the following counts:

1. Truisms. Like the transactions equation, the cash balances equations are truisms. Take any Cambridge equation: Marshall's $P=M/kY$ or Pigou's $P=kR/M$ or Robertson's $P=M/kT$ or Keynes's $p=n/k$, it establishes a proportionate relation between quantity of money and price level.[1]

2. Price Level does not Measure Purchasing Power. Keynes in his *A Treatise on Money* (1930) criticised Pigou's cash balances equation and also his own real balances equation. He pointed out that measuring the price level in wheat, as Piogu did or in terms of consumption units, as Keynes himself did, was a serious defect. The price level in both equations does not measure the purchasing power of money. Measuring the price level in consumption units implies that cash deposits are used only for expenditure on current consumption. But in fact they are held for

"a vast multiplicity of business and personal purposes." By ignoring these aspects, the Cambridge economists have committed a serious mistake.

3. More Importance to Total Deposits. Another defect of the Cambridge equation "lies in its applying to the total deposits considerations which are primarily relevant only to the income deposits." And the importance attached to *k* "is misleading when it is extended beyond the income deposits."

4. Neglects other Factors. Further, the cash balances equation does not tell about changes in the price level due to changes in the proportions in which deposits are held for income, business and savings purposes.

5. Neglect of Saving-Investment Effect. Moreover, it fails to analyse variations in the price level due to saving-investment inequality in the economy.

6. k and Y not Constant. The Cambridge equation, like the transactions equation, assumes *k* and *Y* (or R or T) as constant. This is unrealistic because it is not essential that the cash balances (*k*) and the income of the people (*Y*) should remain constant even during the short period.

7. Fails to Explain Dynamic Behaviour of Prices. The theory argues that changes in the total quantity of money influence the general price level equiproportionally. But the fact is that the quantity of money influences the price level in an essential erratic and unpredictable way. Further, it fails to point out the extent of change in the price level as a result of a given change in the quantity of money in the short period. Thus it fails to explain the dynamic behaviour of prices.

8. Neglects Interest Rate. The cash balances approach is also weak in that it ignores other influences, such as the rate of interest which exerts a decisive and significant influence upon the price level. As pointed out by Keynes, the relation between quantity of money and price level is not direct but indirect via the rate of interest, investment, output, employment and income. This is what the Cambridge equation ignores and hence fails to integrate monetary theory with the theory of value and output.

9. Demand for Money not Interest Inelastic. The neglect of the rate of interest as a causative factor between the quantity of money and the price level led to the assumption that the demand for money is interest inelastic. It means that money performs only the function of medium of exchange and does not possess any utility of its own, such as store of value.

10. Neglect of Goods Market. Further, the omission of the influence of the rate of interest in the cash balances approach led to the failure of neoclassical economists to recognise the interdependence between the commodity and money markets. According to Patinkin, "They laid an undue concentration on the money market a corresponding neglect of the commodity markets, and a resulting 'dehumanising' of the analysis of the effects of monetary changes."

---

1. Here all equations have been changed in terms of the price level.

11. Neglects Real Balance Effect. Patinkin has criticised the Cambridge economists for their failure to integrate the goods market and the money market. This is borne out by the dichotomy which they maintain between the two markets. The dichotomisation implies that the absolute price level in the economy is determined by the demand and supply of money, and the relative price level is determined by the demand and supply of goods. The cash balances approach keeps the two markets rigidly apart. For instance, this approach tells that an increase in the quantity of money leads to an increase in the absolute price level but exercises no influence on the market for goods. This is because of the failure of Cambridge economists to recognise "the real balance effect." The real balance effect shows that a change in the absolute price level does influence the demand and supply of goods. The weakness of cash balances approach lies in ignoring this.

12. Elasticity of Demand for Money not Unity. The cash balances theory establishes that the elasticity of demand for money is unity which implies that the increase in the demand for money leads to a proportionate decrease in the price level. Patinkin holds that "the Cambridge function does not imply uniform elasticity." According to him, this is because of the failure of Cambridge economists to recognise the full implications of the "real balance effect". Patinkin argues that a change in the price level

will cause a real balance effect. For instance, a fall in the price level will increase the real value of cash balances held by the people. So when there is excess demand for money, the demand for goods and services is reduced. In this case, the real balance effect will not cause a proportionate but non-proportionate change in the demand for money. Thus the elasticity of demand for money will not be unity.

13. Neglects Speculative Demand for Money. Another serious weakness of cash balances approach is its failure to consider the speculative demand for money. The neglect of the speculative demand for cash balances makes the demand for money exclusively dependent on money income thereby again neglecting the role of the rate of interest and the store of value function of money.

TRANSACTIONS APPROACH VS. CASH BALANCES APPROACH

There are certain points of similarities between Fisher's transactions approach and the Cambridge cash balances approach. These are discussed as under:

1. Similarities

The two approaches have the following similarities:

1. Same Conclusion. The Fisherian and Cambridge versions lead to the same conclusion that there is a direct and proportional relationship between the quantity of money and the price level and an inverse proportionate relationship between the quantity of money and the value of money.

2. Similar Equations. The two approaches use almost similar equations. Fisher's equation $P = \frac{MV}{T}$ is similar to Robertson's equation $P = \frac{M}{kT}$

However, the only difference is between the two symbols *V* and *k* which

are reciprocal to each other. Whereas $V=\left|\frac{1}{k}\right| \quad k=\left|\frac{1}{V}\right|$. Here *V* refers to the rate of spending and *k* the amount of money which people wish to hold in the form of cash balances or do not want to spend. As these two symbols are reciprocal to each other, the differences in the two equations can be reconciled by substituting $\frac{1}{V}$ for *k* in Robertson's equation and $\frac{1}{k}$ for *V* in Fisher's equation.

3. Money as the Same Phenomenon. The different symbols given to the total quantity of money in the two approaches refer to the same phenomenon. As such *MV+M'V'* of Fisher's equation, *M* of the equations of Pigou and Robertson, and *n* of Keynes' equation refer to the total quantity of money.

2. Dissimilarities

Despite these similarities the two approaches have many dissimilarities:

1. Functions of Money. The two versions emphasize on different functions of money. The Fisherian approach lays emphasis on the medium of exchange function while the Cambridge approach emphasises the store of value function of money.

2. Flow and Stock. In Fisher's approach, money is a flow concept while in the Cambridge approach it is a stock concept. The former relates to a period of time and the latter to a point of time.

3. V and k Different. The meaning given to the two symbols *V* and *k* in the two versions is different. In Fisher's equation *V* refers to the rate of spending and in Robertson's equation *k* refers to the cash balances which people wish to hold. The former emphasises the transactions velocity of circulation and the latter the income velocity.

4. Nature of Price Level. In Fisher's equation, *P* refers to the average price level of all goods and services. But in the cambridge equation *P* refers to the prices of final or consumer goods.

5. Nature of *T*. In Fisher's version, *T* refers to the total amount of goods and services exchanged for money, whereas in the Cambridge version, it refers to the final or consumer goods exchanged for money.

6. Emphasis on Supply and Demand for Money. Fisher's approach emphasises the supply of money, whereas the Cambridge approach emphasises both the demand for money and the supply of money.

7. Different in Nature. The two approaches are different in nature. The Fisherian version is mechanistic because it does not explain how changes in *V* bring about changes in *P*. On the other hand, the Cambridge version is realistic because it studies the psychological factors which influence *k*.

It is on account of these differences that Hansen wrote: "It is not true as is often alleged that the cash balance equation is merely the quantity theory in new algebraic dress."

SUPERIORITY OF CASH BALANCES APPROACH OVER TRANSACTIONS APPROACH

The Cambridge cash balances approach to the quantity theory of money is superior to Fisher's transaction approach in many respects. They are discussed as under:

1. Basis of Liquidity Preference Theory of Interest. The cash balances approach emphasises the importance of holding cash balances rather than the supply of money which is given at a point of time. It thus led Keynes to propound his theory of liquidity preference and of the rate of interest, and to the integration of monetary theory of value and output.

2. Complete Theory. The cash balances version of quantity theory is superior to the transactions version because the former determines the value of money in terms of the demand and supply of money. Thus it is a

complete theory. But in the transactions approach, the determination of value of money is artificially divorced from the theory of value.

3. Discards the Concept of Velocity of Circulation. The cash balances approach is superior to the transactions approach because it discards the concept of the velocity of circulation of money which 'obscures the motives and decisions of people behind it.'

4. Related to the Short Period. Again the cash balances version is more realistic than the transactions version of the quantity theory, because it is related to the short period while the latter is related to the long period. As pointed out by Keynes, "In the long run we may all be dead." So the study of the relationship between quantity of money and price level during the long run is unrealistic.

5. Simple Equations. In the cash balances equations, transactions relating to final goods only are included where *P* refers to the level of final goods. On the other hand, in the transactions equation *P* includes all types of transactions. This creates difficulties in determining the true price level. Thus the former equations are simpler and realistic than the latter.

6. New Formulation in Monetary Theory. Further, the Cambridge equation regards the cash balances held by the people as a function of the level of income. The introduction of income (*Y* or *R* or *T*) in this equation as against *V* (the velocity of circulation of money) in the transaction equation has made the cash balances equation realistic and led to new formulations in monetary theory. "It points out that changes in the level of money income can come about through changes in the price level, through changes in real output or through both at once."

7. Explains Trade Cycles. Hansen regards *k* in the Cambridge equation superior to *V* in Fisher's equation for understanding cyclical fluctuations. According to him, "Drastic and sudden shifts in the desire to hold money, reflected in a change in *k*, may produce large and quickly moving changes in the level of income and prices. In the Cambridge analysis, a shift in *k* may start an upward or downward movement." For instance, when *k* (the fraction of total real income that people wish to hold in cash balances) increases because of low business expectations, the price level falls, and

vice versa.

8. Study of Subjective Factors. As a corollary to the above, *V* in Fisher's equation is mechanistic while *k* in the Cambridge equation is realistic. The subjective factors behind variations in *k* have led to the study of such factors as expectations, uncertainty, motives for liquidity, and the rate of interest in modern monetary theory. In this sense, it can be justifiably said that, "the Cambridge equation moves us on from the tautology represented by the equation of exchange to a study of economic behaviour."

9. Applicable under All Circumstances. Fisher's transactions approach holds true only under full employment. But the cash balances approach holds under all circumstances whether there is full employment or less than full employment.

10. Based on Micro Factors. The Cambridge version is superior to the Fisherian version because it is based on micro factors like individual decisions and behaviours. On the other hand, the Fisherian version is based on macro factors like *T*, total velocity of circulation, etc.

EXERCISES

1. Critically examine Fisher's Quantity Theory of Money.
2. Discuss the Cambridge Cash Balances Approach to the Quantity Theory of Money. How far is it superior to the Cash Transactions Approach ?
3. Give the similarities and dissimilarities of the Quantity Theory of Money and Cambridge cash balances theory.
4. Write notes on.

   (a) Fisher's Equation of Exchange.

   (b) Keynes' Real Balance Equation.

5. What do you mean by value of money ? What determines value of money, according to Fisher ?

CHAPTER 26

THE KEYNESIAN THEORY OF MONEY AND PRICES

INTRODUCTION

In Chapter 21 of his *General Theory*, Keynes attacked the classical quantity theorists for keeping separate monetary theory and value theory. He then presented a reformulated quantity theory of money which brought about a transition from a monetary theory of prices to a monetary theory of output. In doing this, Keynes made an attempt to intergrate monetary theory with value theory and also linked the theory of interest into monetary theory. But "it is through the theory of output that value theory and monetary theory are brought into just a position with each other."[1]

Keynes does not agree with the older quantity theorists that there is a direct and proportional relationship between quantity of money and prices. According to him, the effect of a change in the quantity of money on prices is indirect and non-proportional.

Keynes complains "that economics has been divided into two compartments with no doors or windows between the theory of value and the theory of money and prices."[2] This dichotomy between the *relative price level* (as determined by demand and supply of goods) and the *absolute price level* (as determined by demand and supply of money) arises from the failure of the classical monetary economists to integrate value theory with monetary theory. Consequently, changes in the money supply affect only the absolute price level but exercise no influence on the relative price level.

1. D. Dillard, *op. cit.*, pp. 224-225.

Further, Keynes criticises the classical theory of static equilibrium in which money is regarded as neutral and does not influence the economy's real equilibrium relating to relative prices. According to him, the problems of the real world are related to the theory of shifting equilibrium whereas money enters as a "link between the present and future".

KEYNES'S REFORMULATED QUANTITY THEORY OF MONEY

In his reformulated quantity theory of money, Keynes shows the relationship between money, output and prices. The theory points out that (a) so long as there is unemployment, output will change in the same proportion as the quantity of money, and (b) when there is full employment, prices will change in the same proportion as the quantity of money.

Assumptions

This theory is based on the following assumptions:

1. All factors of production are in perfectly elastic supply so long as there is any unemployment.

2. All unemployed factors are homogeneous, perfectly divisible and interchangeable.

3. There are constant returns to scale so that prices do not rise or fall as output increases.

4. Effective demand and quantity of money change in the same proportion so long as there are any unemployed resources.

Explanation

Given these assumptions, the Keynesian chain of causation between changes in the quantity of money and in prices is an *indirect* one through

the rate of interest. So when the quantity of money is increased, its first impact is on the rate of interest which tends to fall. Given the marginal efficiency of capital, a fall in the rate of interest will increase the volume of investment. The increased investment will raise effective demand through the multiplier effect thereby increasing income, output and employment. Since the supply curve of factors of production is perfectly elastic in a situation of unemployment, wage and non-wage factors are available at constant rate of remuneration. There being constant returns to scale, prices do not rise with the increase in output so long as there is any unemployment. Under the circumstances, output and employment will increase in the same proportion as effective demand, and the effective demand will increase in the same proportion as the quantity of money. Thus so long as there is unemployment, output will change in the same proportion as the quantity of money, and there will be no change in prices; and when there is full employment, prices will change in the same proportion as the quantity of money. Therefore, the reformulated quantity theory of money stresses the point that with increase in the quantity of money, prices rise only when the level of full employment is reached, and not before this.

---

2. A.H. Hansen, *op. cit.*, p. 184.

This reformulated quantity theory of money is illustrated in Figure 1 (A) and (B) where *OTC* is the output curve relating to the quantity of money and *PRC* is the price curve relating to the quantity of money. Panel A of the figure shows that as the quantity of money increases from *O* to *M*, the level of output also rises along the *OT* portion of the *OTC* curve. As the quantity of money reaches *OM* level, full employment output OQ_F is being produced. But after point *T* the output curve becomes vertical because any further increase in the quantity of money cannot raise output beyond the full employment level OQ_F.

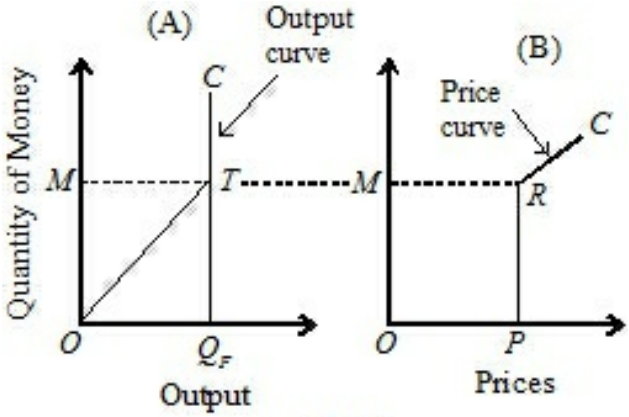


Fig. 1.

Panel B of the figure shows the relationship between quantity of money and prices. So long as there is unemployment, prices remain constant whatever the increase in the quantity of money. Prices start rising only after the full employment level is reached. In the figure, the price level *OP* remains constant at the *OM* quantity of money corresponding to the full employment level of output OQ_F. But an increase in the quantity of money above *OM* raises prices in the same proportion as the quantity of money. This is shown by the *RC* portion of the price curve *PRC*.

The Complicated Model. Keynes himself pointed out that the real world is so complicated that the simplifying assumptions upon which the reformulated quantity theory of money is based, will not hold. According to him, the following possible complications would qualify the statement that so long as there is unemployment, employment will change in the same proportion as the quantity of money, and when there is full employment, *prices* will change in the same proportion as the quantity of money."

(1) "Effective demand will not change in exact proportion to the quantity of money.

(2) Since resources are homogenous, there will be diminishing, and not constant returns as employment gradually increases.

(3) Since resources are not interchangeable, some commodities will reach a condition of inelastic supply while there are still unemployed resources available for the production of other commodities.

(4) The wage-unit will tend to rise, before full employment has been reached.

(5) The remunerations of factors entering into marginal cost will not all change in the same proportion."[3]

Taking into account these complications, it is clear that the reformulated quantity theory of money does not hold. An increase in effective demand will not change in exact proportion to the quantity of money, but it will partly spend itself in increasing output and partly in increasing the price

level. So long as there are unemployed resources, the general price level will not rise much as output increases. But a sudden large increase in aggregate demand will encounter bottlenecks when resources are still unemployed. It may be that the supply of some factors becomes inelastic or others may be in short supply and are not interchangeable. This may lead to increase in marginal cost and price. Price would accordingly rise above average unit cost and profits would increase rapidly which, in turn, tend to raise money wages owing to trade union pressures. Diminishing returns may also set in. As full employment is reached, the elasticity of supply of output falls to zero[4] and prices rise in proportion to the increase in the quantity of money. The chain of causation between changes in money (*M*) and prices (*P*) can be shown thus:

---

3. J.M. Keynes, *op. cit.*, pp. 296-97.

$$M\uparrow \rightarrow R\downarrow \rightarrow I\uparrow \rightarrow Q\uparrow \rightarrow C\uparrow \rightarrow P\uparrow$$

(*R*) Rate of Interest, (*I*) Investment, (*Q*) Output, (*C*) Costs

The complicated model of the Keynesian theory of money and prices is shown diagrammatically in Figure 2 in terms of aggregate supply (S) and aggregate demand (D) curves. The price level is measured on the vertical axis and output on the horizontal axis.

According to Keynes, an increase in the quantity of money increases aggregate money demand on investment as a result of the fall in the rate of interest. This increases output and employment in the beginning but not the price level. In the figure, the increase in the aggregate money demand from D_1 to D_2 raises output from OQ_1 to OQ_2 but the price level remains constant at OP. As aggregate money demand increases further from D_2 to D_3, output increases

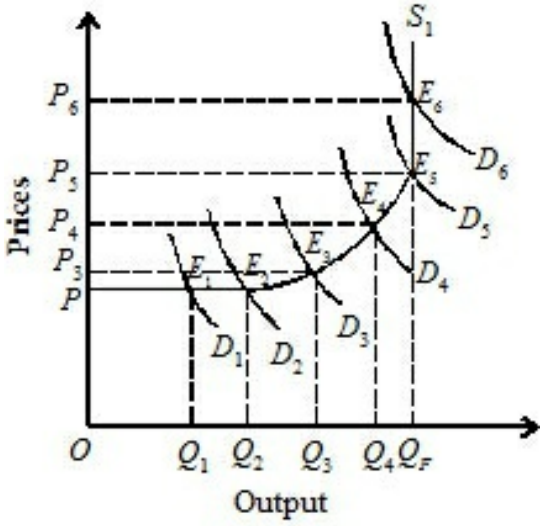


Fig. 2.

from OQ_2 to OQ_3, and the price level also rises to OP_3. This is because costs rise as bottlenecks develop through the immobility of resources. Diminishing returns set in and less efficient labour and capital are employed. Output increases at a slower rate than a given increase in aggregate *money demand*, and this leads to higher prices. As full employment is approached, bottlenecks increase. Further, rising prices lead to increased demand, especially for stocks. Thus prices rise at an increasing rate. This is shown over the range E_3 E_5 in the figure. But when the economy reaches the full employment level of output, any further increase in aggregate money demand brings about a proportionate increase in the price level but output remains unchanged at that level. This is shown in the figure when the demand curve D_5 shifts upward to D_6 and the price level increases from OP_5 to OP_6 while the level of output remains constant at OQ_F.

SUPERIORITY OF THE KEYNESIAN THEORY OVER THE TRADITIONAL QUANTITY THEORY OF MONEY

The Keynesian theory of money and prices is surperior to the traditional quantity theory of money for the following reasons.

Keynes's reformulated quantity theory of money is superior to the traditional approach in that he discards the old view that the relationship between quantity of money and prices is direct and proportional. Instead, he establishes an indirect and non-proportional relationship between quantity of money and prices.

---

4. The elasticity of output with respect to demand $e = DdO/OdD$

In establishing such a relationship, Keynes brought about a transition from a pure monetary theory of prices to a monetary theory of output and employment. In so doing, he integrates monetary theory with value theory. He and also with the theory of output and employment through the rate of interest. In fact, the integration between monetary theory and value theory is done through the theory of output in which the rate of interest plays the

crucial role. The Keynesain theory is, therefore, superior to the traditional quantity theory of money because it does not keep the real and monetary sectors of the economy into two separate compartments with "no doors or windows between the theory of value and the theory of money and prices."

Again, the traditional quantity theory is based on the unrealistic asumption of full employment of resources. Under this assumption, a given increase in the quantity of money always leads to a proportionate increase in the price level. Keynes, on the other hand, believes that full employment is an exception. Therefore, so long as there is unemployment, output and employment will change in the same proportion as the quantity of money, but there willl be no change in prices. When there is full employment, prices will change in the same proportion as the quantity of money. Thus the Keynesian analysis is superior to the traditional analysis because it studies the relationship between quantity of money and prices both under unemployment and full employment situations.

Further, the Keynesian theory is superior to the traditional quantity theory of money in that it emphasises important policy implications. The traditional theory believes that every increase in the quantity of money leads to inflation. Keynes, on the other hand, establishes that so long as there is unemployment, the rise in prices is gradual and there is no danger of inflation. It is only when the economy reaches the level of full employment that the rise in prices is inflationary with every increase in the quantity of money. Thus "this approach has the virtue of emphasising that the objectives of full employment and price stability may be inherently irreconcilable."

CRITICISMS OF KEYNES' THEORY OF MONEY AND PRICES

Keynes' views on money and prices have been criticised by the monetarists on the following grounds:

1. Direct Relation. Keynes mistakenly took prices as fixed so that the effect of money appears in his analysis in terms of quantity of goods traded rather than their average prices. That is why Keynes adopted an

indirect mechanism through bond prices, interest rates and investment of the effects of monetary changes on economic activity. But the actual effects of monetary changes are direct rather than indirect.

2. Stable Demand for Money. Keynes assumed that monetary changes are largely absorbed by changes in the demand for money. But Friedman has shown on the basis of his empirical studies that the demand for money is highly stable.

3. Nature of Money. Keynes failed to understand the true nature of money. He believed that money could be exchanged for bonds only. In fact, money can be exchanged for many different types of assets like bonds, securities, physical assets, human wealth, etc.

4. Effect of Money. Since Keynes wrote for a depression period, this led him to conclude that money had little effect on income. According to Friedman, it was the contraction of money that precipitated the depression. It was, therefore, wrong on the part of Keynes to argue that money had little effect on income. Money does affect national income.

EXERCISES

1. Show how the Theory of Value has been integrated with the Theory of Money by Keynes.
2. How has Keynes correlated money and rate of interest on the one hand and rate of interest, employment and prices on the other hand ?
3. Critically examine the Keynesian theory of money and prices.
4. Explain Keynes' theory of money and prices. In what ways it is superior to the Quantity Theory of Money?

CHAPTER 27

FRIEDMAN'S RESTATEMENT OF THE QUANTITY THEORY OF MONEY

INTRODUCTION

Following the publication of Keynes's *General Theory* in 1936, economists discarded the traditional quantity theory of money. But at the University of Chicago, "the quantity theory continued to be a central and vigorous part of the oral tradition throughout the 1930s and 1940s." At Chicago, Milton Friedman, Henry Simons, Lloyd Mints, Frank Knight and Jacob Viner taught and developed 'a more subtle and relevant version' of the quantity theory of money in its theoretical form "in which the quantity theory was connected and integrated with general price theory." The foremost exponent of the Chicago version of the quantity theory of money who led to the so-called "Monetarist Revolution" is Professor Friedman. He, in his essay "The Quantity Theory of Money—A Restatement" published in 1956', set down a particular model of quantity theory of money. This is discussed below.

1. Milton Friedman (ed.), *Studies in the Quantity Theory of Money*, 1956, pp. 3-21.

FRIEDMAN'S THEORY

In his reformulation of the quantity theory, Friedman asserts that "the quantity theory is in the first instance a theory of the demand for money. It is not a theory of output, or of money income, or of the price level." The demand for money on the part of ultimate wealth holders is formally

identical with that of the demand for a consumption service. He regards the amount of real cash balances (M/P) as a commodity which is demanded because it yields services to the person who holds it. Thus money is an asset or capital good. Hence the demand for money forms part of capital or wealth theory. For ultimate wealth holders, the demand for money, in real terms, may be expected to be a function primarily of the following variables:

1. Total Wealth. The total wealth is the analogue of the budget constraint. It is the total that must be divided among various forms of assets. In practice, estimates of total wealth are seldom available. Instead, income may serve as an index of wealth. Thus, according to Friedman, income is a surrogate of wealth.

2. The Division of Wealth between Human and Non-Human Forms. The major source of wealth is the productive capacity of human beings which is human wealth. But the conversion of human wealth into non-human wealth or the reverse is subject to institutional constraints. This can be done by using current earnings to purchase non-human wealth or by using non-human wealth to finance the acquisition of skills. Thus the fraction of total wealth in the form of non-human wealth is an additional important variable. Friedman calls the ratio of non-human to human wealth or the ratio of wealth to income as *w*.

3. The Expected Rates of Return on Money and Other Assets. These rates of return are the counterparts of the prices of a commodity and its substitutes and complements in the theory of consumer demand. The nominal rate of return may be *zero* as it generally is on currency, or *negative* as it sometimes is on demand deposits, subject to net service charges, or *positive* as it is on demand deposits on which interest is paid, and generally on time deposits. The nominal rate of return on other assets consists of two parts: *first*, any currently paid yield or cost, such as interest on bonds, dividends on equities, and costs of storage on physical assets, and *second*, changes in the prices of these assets which become especially important under conditions of inflation or deflation.

4. Other Variables. Variables other than income may affect the utility attached to the services of money which determine liquidity proper.

Besides liquidity, variables are the tastes and preferences of wealth holders. Another variable is trading in existing capital goods by ultimate wealth holders. These variables also determine the demand function for money alongwith other forms of wealth. Such variables are noted as *u* by Friedman.

Broadly, total wealth includes all sources of income or consumable services. It is capitalised income. By income, Friedman means "permanent income" which is the average expected yield on wealth during its life time.

Forms of Wealth. Wealth can be held in five different forms: money, bonds, equities, physical goods, and human capital. Each form of wealth has a unique characteristic of its own and a different yield.

1. *Money* is taken in the broadest sense to include currency, demand deposits and time deposits which yield interest on deposits. Thus money is luxury good. It also yields real return in the form of convenience, security, etc. to the holder which is measured in terms of the general price level (*P*).

2. *Bonds* are defined as claim to a time stream of payments that are fixed in *nominal units.*

3. *Equities* are defined as a claim to a time stream of payments that are fixed in *real* units.

4. *Physical goods or non-human goods* are inventories of producer and consumer durables.

5. *Human capital* is the productive capacity of human beings.

Thus each form of wealth has a unique characteristic of its own and a different yield either explicitly in the form of interest, dividends, labour income, etc., or implicitly in the form of services of money measured in terms of *P*, and inventories. The present discounted value of these expected income flows from these five forms of wealth constitutes the current value of wealth which can be expressed as:

$$W = \frac{y}{r}$$

where W is the current value of total wealth, y is the total flow of expected income from the five forms of wealth, and r is the interest rate. This equation shows that wealth is capitalised income. Friedman in his latest empirical study *Monetary Trends in the United States and the United Kingdom* (1982) gives the following demand function for money for an individual wealth holder with slightly different notations from his original study of 1956 as:

$$M/P = f\,(y, w; R_m, R_b, R_e, g_p, u)$$

where M is the total stock of money demanded; P is the price level; y is the real income; w is the fraction of wealth in non-human form; R_m is the expected nominal rate of return on money; R_b is the expected rate of return on bonds, including expected changes in their prices; R_e is the expected nominal rate of return on equities, including expected changes in their prices; $g_p = (1/P)\,(dP/dt)$ is the expected rate of change of prices of goods and hence the expected nominal rate of return on physical assets; and u stands for variables other than income that may affect the utility attached to the services of money.

The demand function for business is roughly similar, although the division of total wealth and human wealth is not very useful since a firm can buy and sell in the market place and hire its human wealth at will. But the other factors are important.

The aggregate demand function for money is the summation of individual demand functions with M and y referring to per capita money holdings and per capita real income respectivley, and w is the fraction of aggregate wealth in non-human form.

The demand function for money leads to the conclusion that a rise in expected yields on different assets (R_b, R_e and g_p) reduces the amount of money demanded by a wealth holder, and that an increase in wealth raises the demand for money. The income to which cash balances (M/P) are adjusted is the expected long-term level of income rather than current income being received. Empirical evidence suggests that the income elasticity of demand for money is greater than unity which means that

income velocity is falling over the long run. This means that the long-run demand for money function is stable and is relatively interest inelastic, as shown in Fig. 1 where M_D is the demand for money curve. If there is change in the interest rate, the long-run demand for money is negligible.

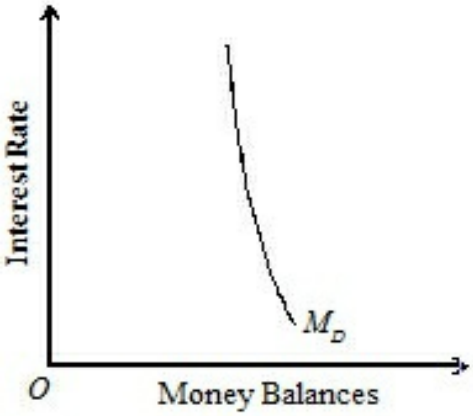


Fig. 1

In Friedman's restatement of the quantity theory of money, the supply of money is independent of the demand for money. The supply of money is unstable due to the actions of monetary authorities. On the other hand, the demand for money is stable. It means that money which people want to hold in cash or bank deposits is related in a fixed way to their permanent income. If the central bank increases the supply of money by purchasing securities, people who sell securities will find their holdings of money have increased in relation to their permanent income. They will, therefore, spend their excess holdings of money, partly on assets and partly on consumer goods and services. This spending will reduce their money balances and at the same time raise the nominal income. On the contrary, a reduction in the money supply by selling securities on the part of the central bank will reduce the holdings of money of the buyers of securities in relation to their permanent income. They will, therefore, raise their money holdings partly by selling their assets and partly by reducing their consumption expenditure on goods and services. This will tend to reduce nominal income. Thus, on both counts, the demand for money remains stable. According to Friedman, a change in the supply of money causes a proportionate change in the price level or income or in both. Given the

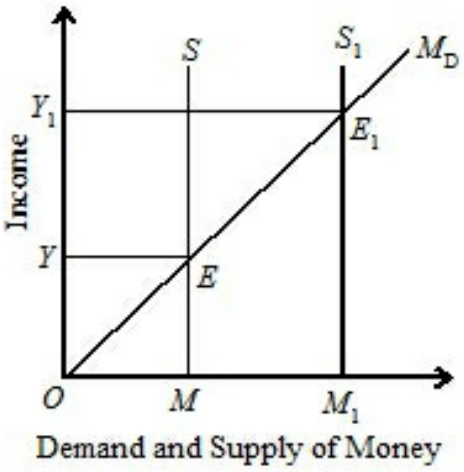


Fig. 2.

demand for money, it is possible to predict the effects of changes in the supply of money on total expenditure and income. If the economy is operating at less than full employment level, an increase in the supply of money will raise output and employment with a rise in total expenditure. But this is only possible in the short run. Friedman's quantity theory of moeny is explained in Figure 2, where income (Y) is measured on the vertical axis and the demand for and supply of money are measured on the horizontal axis. M_D is the demand for money curve which varies with income. MS is the money supply curve which is perfectly inelalstic to changes in income. The two curves intersect at E and determine the equilibrium income OY. If the money supply rises, the MS curve shifts to the right to M_1S_1. As a result, the money supply is greater than the demand for money which raises total expenditure until new equilibrium is established at E_1 between M_D and M_1S_1 curves and income rises to OY_1.

Thus Friedman presents the quantity theory as the theory of the demand for money and the demand for money is assumed to depend on asset prices or relative returns and wealth or income. He shows how a theory of the stable demand for money becomes a theory of prices and output. A discrepancy between the nominal quantity of money demanded and the nominal quantity of money supplied will be evident primarily in attempted spending. As the demand for money changes in response to changes in its determinants, it follows that substantial changes in prices or nominal income are almost invariably the result of changes in the nominal supply of money.

Its Criticisms

Friedman's reformulation of the quantity theory of money has evoked much controversy and has led to empirical verification on the part of the Keynesians and the Monetarists. Some of the criticisms levelled against the theory are discussed as under:

1. Very Broad Definition of Money. Friedman has been criticised for using the broad definition of money which not only includes currency and demand deposits (M_1) but also time deposits with commercial banks (M_2). This broad definition leads to the obvious conclusion that the interest

elasticity of the demand for money is negligible. If the rate of interest increases on time deposits, the demand for them (M_2) rises. But the demand for currency and demand deposits (M_1) falls. So the overall effect of the rate of interest will be negligible on the demand for money. But Friedman's analysis is weak in that he does not make a choice between long-term and short-term interest rates. In fact, if demand deposits (M_1) are used, a short-term rate is preferable, while a long-term rate is better with time deposits (M_2). Such an interest rate structure is bound to influence the demand for money.

2. Money not a Luxury Good. Friedman regards money as a luxury good because of the inclusion of time deposits in money. This is based on his finding that there is higher trend rate of the money supply than income in the United States. But no such 'luxury effect' has been found in the case of England.

3. More Importance to Wealth Variables. In Friedman's demand for money function, wealth variables are preferable to income and the operation of wealth and income variables simultaneously does not seem to be justified. As pointed out by Johnson, income is the return on wealth, and wealth is the present value of income. The presence of the rate of interest and one of these variables in the demand for money function would appear to make the other superfluous.

4. Money Supply not Exogenous. Friedman takes the supply of money to be unstable. The supply of money is varied by the monetary authorities in an exogenous manner in Friedman's system. But the fact is that in the United States, the money supply consists of bank deposits created by changes in bank lending. Bank lending, in turn, is based upon bank reserves which expand and contract with (a) deposits and withdrawals of currency by non-bank financial intermediaries; (b) borrowings by commercial banks from the Federal Reserve System; (c) inflows and outflows of money from and to abroad; and (d) purchase and sale of securities by the Federal Reserve System. The first three items definitely impart an endogenous element to the money supply. Thus the money supply is not exclusively exogenous, as assumed by Friedman. It is mostly endogenous.

5. Ignores the Effect of Other Variables on Money Supply. Friedman also ignores the effect of prices, output or interest rates on the money supply. But there is considerable empirical evidence that the money supply can be expressed as a function of the above variables.

6. Does not consider Time Factor. Friedman does not tell about the timing and speed of adjustment or the length of time to which his theory applies.

7. No Positive Correlation between Money Supply and Money GNP. Money supply and money GNP have been found to be positively correlated in Friedman's findings. But, according to Kaldor,[2] in Britain the best correlation is to be found between the quarterly variations in the amount of cash held in the form of notes and coins by the public and corresponding variations in personal consumption at market prices, and not between money supply and GNP.

---

2. N. Kaldor, 'The New Monetarism,' *L.B.R.*, July 1970

Conclusion. Despite these criticisms, Johnson writes: "Friedman's application to monetary theory of the basic principle of capital theory—that is the yield on capital, and capital the present value of income—is probably the most important development in monetary theory since Keynes's *General Theory*. Its theoretical significance lies in the conceptual integration of wealth and income as influences on behaviour."

FRIEDMAN VS KEYNES

Friedman's demand for money function differs from that of Keynes' in many ways which are discussed as under.

First, Friedman uses a broader definition of money than that of Keynes in order to explain his demand for money function. He treats money as an asset or capital good capable of serving as a temporary abode of purchasing power. It is held for the stream of income or consumable services which it renders. On the other hand, the Keynesian definition of money consists of demand deposits and non-interest bearing debt of the

government.

Second, Friedman postulates a demand for money function quite different from that of Keynes. The demand for money on the part of wealth holders is a function of many variables. These are $R_{\mathbf{m}}$, the yield on money; $R_{\mathbf{b}}$, the yield on bonds; $R_{\mathbf{e}}$, the yield on securities; $g_{\mathbf{p}}$, the yield on physical assets; and *u* referring to other variables. In the Keynesian theory, the demand for money as an asset is confined to just bonds where interest rates are the relevant cost of holding money.

Third, there is also the difference between the monetary mechanisms of Keynes and Friedman as to how changes in the quantity of money affect economic activity. According to Keynes, monetary changes affect economic activity *indirectly* through bond prices and interest rates. The monetary authorities increase the money supply by purchasing bonds which raises their prices and reduces the yield on them. Lower yield on bonds induces people to put their money elsewhere, such as investment in new productive capital that will increase output and income. On the other hand, in Friedman's theory monetary disturbances will *directly* affect prices and production of all types of goods since people will buy or sell any asset held by them. Friedman emphasises that the market interest rates play only a small part of the total spectrum of rates that are relevant.

Fourth, there is the difference between the two approaches with regard to the motives for holding money balances. Keynes divides money balances into "active" and "idle" categories. The former consist of transactions and precautionary motives, and the latter consist of the speculative motive for holding money. On the other hand, Friedman makes no such division of money balances. According to him, money is held for a variety of different purposes which determine the total volume of assets held such as money, physical assets, total wealth, human wealth, and general preferences, tastes and anticipations.

Fifth, in his analysis, Friedman introduces permanent income and nominal income to explain his theory. Permanent income is the amount a wealth holder can consume while maintaining his wealth intact. Nominal income is measured in the prevailing units of currency. It depends on both prices

and quantities of goods traded. Keynes, on the other hand, does not make such a distinction.

EXERCISES

1. Explain the contribution of Milton Friedman to the Quantity Theory of Money.
2. Critically examine the statement that the revised version of the quantity theory of money is that there exists a stable demand function for real money holdings.
3. Explain Friedman's reformulation of the Quantity Theory of Money. How does it differ from that of Keynes?
4. Comment on the following statements:

   (a) The Quantity Theory of Money is a Theory of Demand for Money.

   (b) The Quantity Theory of Money is a Theory of Money-Income Determination.

CHAPTER 28

THE SUPPLY OF MONEY

DEFINITIONS OF MONEY SUPPLY

The supply of money is a stock at a particular point of time, though it conveys the idea of a flow over time. The term 'the supply of money' is synonymous with such terms as 'money stock', 'stock of money', 'money supply' and 'quantity of money'. The supply of money at any moment is the total amount of money in the economy. There are three alternative views regarding the definition or measures of money supply. The most common view is associated with the traditional and Keynesian thinking which stresses the medium of exchange function of money. According to this view, money supply is defined as currency with the public and demand deposits with commercial banks. Demand deposits are savings and current accounts of depositors in a commercial bank. They are the liquid form of money because depositors can draw cheques for any amount lying in their accounts and the bank has to make immediate payment on demand. Demand deposits with commercial banks plus currency with the public are together denoted as M_1, the money supply. This is regarded as a narrower definition of the money supply.

The second definition is broader and is associated with the modern quantity theorists headed by Friedman. Friedman defines the money supply at any moment of time as "literally the number of dollars people are carrying around in their pockets, the number of dollars they have to their credit at banks or dollars they have to their credit at banks in the form of demand deposits, and also commercial bank time deposits." Time deposits are fixed deposits of customers in a commercial bank. Such

deposits earn a fixed rate of interest varying with the time period for which the amount is deposited. Money can be withdrawn before the expiry of that period by paying a penal rate of interest to the bank. So time deposits possess liquidity and are included in the money supply by Friedman. Thus this definition includes M_1 plus time deposits of commercial banks in the supply of money. This wider definition is characterised as M_2 in America and M_3 in Britain and India. It stresses the store of value function of money or what Friedman says, 'a temporary abode of purchasing power'.

The third definition is the broadest and is associated with Gurley and Shaw.[1] They include in the supply of money, M_2 plus deposits of savings banks, building societies, loan associations, and deposits of other credit and financial institutions.

The choice between these alternative definitions of the money supply depends on two considerations: *One*, "a particular choice of definition may facilitate or blur the analysis of the various motives for holding cash;"[2] and *two* from the point of view of monetary policy an appropriate definition should include the area over which the monetary authorities can have direct influence. If these two criteria are applied, none of the three definitions is wholly satisfactory.

The first definition of money supply may be analytically better because M_1 is a sure medium of exchange. But M_1 is an inferior store of value because it earns no rate of interest, as is earned by time deposits. Further, the central bank can have control over a narrower area if only demand deposits are included in the money supply.

The second definition that includes time deposits (M_2) in the supply of money is less satisfactory analytically because "in a highly developed financial structure, it is important to consider separately the motives for holding means of payment and time deposits." Unlike demand deposits, time deposits are not a perfect liquid form of money. This is because the amount lying in them can be withdrawn immediately by cheques. Normally, it cannot be withdrawn before the due date of expiry of deposit. In case a depositor wants his money earlier, he has to give a notice to the

bank which allows the withdrawal after charging a penal interest rate from the depositor. Thus time deposits lack perfect liquidity and cannot be included in the money supply. But this definition is more appropriate from the point of view of monetary policy because the central bank can exercise control over a wider area that includes both demand and time deposits held by commercial banks.

The third definition of money supply that includes M_2 plus deposits of non-bank financial institutions is unsatisfactory on both the criteria. Firstly, they do not serve the medium of exchange function of money. Secondly, they almost remain outside the area of control of the central bank. The only advantage they possess is that they are highly liquid store of value. Despite this merit, deposits of non-bank financial institutions are not included in the definition of money supply.

---

1. *J.G. Gurley and E.S. Shaw,* Money in a Theory of Finance, *2/e, 1965.*

2. *R.J. Ball,* Inflation and the Theory of Money, *1964.*

DETERMINANTS OF MONEY SUPPLY

There are two theories of the determination of the money supply. According to the *first view,* the money supply is determined exogenously by the central bank. The *second view* holds that the money supply is determined endogenously by changes in the economic activity which affect people's desire to hold currency relative to deposits, the rate of interest, etc.

Thus the determinants of money supply are both exogenous and endogenous which can be described broadly as: the minimum cash reserve ratio, the level of bank reserves, and the desire of the people to hold currency relative to deposits. The last two determinants together are called the *monetary base* or *high powered money*.

1. The Required Reserve Ratio

The required reserve ratio (or the minimum cash reserve ratio or the

reserve deposit ratio) is an important determinant of the money supply. An increase in the required reserve ratio reduces the supply of money with commercial banks and a decrease in required reserve ratio increases the money supply. The *RRr* is the ratio of cash to current and time deposit liabilities which is determined by law. Every commercial bank is required to keep a certain percentage of these liabilities in the form of deposits with the central bank of the country. But notes or cash held by commercial banks in their tills are not included in the minimum required reserve ratio.

But the short-term assets along with cash are regarded as the liquid assets of a commercial bank. In India the statutory liquidity ratio (SLR) has been fixed by law as an additional measure to determine the money supply. The SLR is called secondary reserve ratio in other countries while the required reserve ratio is referred to as the primary ratio. The raising of the SLR has the effect of reducing the money supply with commercial banks for lending purposes, and the lowering of the SLR tends in increase the money supply with banks for advances.

2. The Level of Bank Reserves

The level of bank reserves is another determinant of the money supply. Commercial bank reserves consist of reserves on deposits with the central bank and currency in their tills or vaults. It is the central bank of the country that influences the reserves of commercial banks in order to determine the supply of money. The central bank requires all commercial banks to hold reserves equal to a fixed percentage of both time and demand deposits. These are legal minimum or required reserves. Required reserves (*RR*) are determined by the required reserve ratio (*RRr*) and the level of deposits (D) of a commercial bank: $RR=RRr\times D$. If deposits amount to Rs 80 lakhs and required reserve ratio is 20 percent, then the required reserves will be 20% × 80=Rs 16 lakhs. If the reserve ratio is reduced to 10 per cent, the required reserves will also be reduced to Rs 8 lakhs. Thus the higher the reserve ratio, the higher the required reserves to be kept by a bank, and vice versa. But it is the excess reserves (*ER*) which are important for the determination of the money supply. Excess reserves are the difference between total reserves (*TR*) and required reserves (*RR*): ER=TR-RR. If total reserves are Rs 80 lakhs and required reserves are Rs

16 lakhs, then the excess reserves are Rs 64 lakhs (Rs 80 - 16 lakhs). When required reserves are reduced to Rs 8 lakhs, the excess reserves increase to Rs 72 lakhs. It is the excess reserves of a commercial bank which influence the size of its deposit liabilities. A commercial bank advances loans equal to its excess reserves which are an important component of the money supply. To determine the supply of money with a commercial bank, the central bank influences its reserves by adopting open market operations and discount rate policy.

Open market operations refer to the purchase and sale of government securities and other types of assets like bills, securities, bonds, etc., both government and private in the open market. When the central bank buys or sells securities in the open market, the level of bank reserves expands or contracts. The purchase of securities by the central bank is paid for with cheques to the holders of securities who, in turn, deposit them in commercial banks thereby increasing the level of bank reserves. The opposite is the case when the central bank sells securities to the public and banks who make payments to the central bank through cash and cheques thereby reducing the level of bank reserves.

The discount rate policy affects the money supply by influencing the cost and supply of bank credit to commercial banks. The discount rate, known as the bank rate in India, is the interest rate at which commercial banks borrow from the central bank. A high discount rate means that commercial banks get less amount by selling securities to the central bank. The commercial banks, in turn, raise their lending rates to the public thereby making advances dearer for them. Thus there will be contraction of credit and the level of commercial bank reserves. Opposite is the case when the bank rate is lowered. It tends to expand credit and consequently bank reserves.

It should be noted that commercial bank reserves are affected significantly only when open market operations and discount rate policy supplement each other. Otherwise, their effectiveness as determinants of bank reserves and consequently of money supply is limited.

3. Public's Desire to Hold Currency and Deposits

People's desire to hold currency (or cash) relative to deposits in commercial banks also determines the money supply. If people are in the habit of keeping less in cash and more in deposits with the commercial banks, the money supply will be large. This is because banks can create more money with larger deposits. On the contrary, if people do not have banking habits and prefer to keep their money holdings in cash, credit creation by banks will be less and the money supply will be at a low level.

4. High-Powered Money

The current practice is to explain the determinants of money supply in terms of the monetary base or high-powered money. High-powered money is the sum of commercial bank reserves and currency (notes and coins) held by the public. High-powered money is the base for the expansion of bank deposits and creation of the money supply. The supply of money varies directly with changes in the monetary base, and inversely with the currency and reserve ratios.

5. Other Factors

Money supply is a function not only of the high-powered money determined by the monetary authorities, but of interest rates, income and other factors. The latter factors change the proportion of money balances that the public holds as cash. Changes in business activity can change the behaviour of banks and the public and thus affect the money supply. Hence the money supply is not only an exogenous controllable item but also an endogenously determined item.

Conclusion. We have discussed above the factors which determine money supply through the creation of bank credit. But money supply and bank credit are indirectly related to each other. When the money supply increases, a part of it is saved in banks depending upon the depositors' propensity to save. These savings become deposits of commercial banks who, in turn, lend after meeting the statutory reserve requirements. Thus with every increase in the money supply, the bank credit goes up. But it may not happen in exactly the same proportion due to the following factors:

(a) The marginal propensity to save does not remain constant. It varies from time to time depending on changes in income levels, prices, and subjective factors.

(b) Banks may also create more or less credit due to the operation of leakages in the credit creation process.

(c) The velocity of circulation of money also affects the money supply. If the velocity of money circulation increases, the bank credit may nor fall even after a decrease in the money supply. The central bank has little control over the velocity of money which may adversely affect bank credit.

HIGH-POWERED MONEY AND THE MONEY MULTIPLIER

The current practice is to explain the determinants of money supply in terms of the monetary base or high-powered money. High -powered money is the sum of commercial bank reserves and currency (notes and coins) held by the public. High-powered money is the base for the expansion of bank deposits and creation of the money supply. The supply of money varies directly with changes in the monetary base, and inversely with the currency and reserve ratios.

The use of high-powered money consists of the demand of commercial banks for the legal limit or required reserves with the central bank and excess reserves and the demand of the public for currency. Thus high-powered money $H=C+RR+ER$ where C represents currency, RR the required reserves and ER the excess reserves.

A commercial bank's required reserves depend upon its deposits. But a bank usually holds reserves in excess of its required reserves. In fact, banks do not advance loans up to the legal limits but precisely less than that. This is to meet unanticipated cash withdrawals or adverse clearing balances. Hence the need arises for maintaining excess reserves by them. The money supply is thus determined by the required reserve ratio and the excess reserve ratio of commercial banks. The required reserve ration (RRr) is the ratio of required reserves to to deposits (RR/D), and the

excess reserve ratio (*ERr*) is the ratio or excess reserves to deposits (*ER/D*).

Currency held by the public is another component of high-powered money. The demand for currency by the public is expressed as a proportion of bank deposits. Thus the currency ratio *Cr=C/D*, where *C* is the currency and *D* deposits. The currency ratio is influenced by such factors as changes in income levels of the people, the use of credit instruments by the public, and uncertainties in economic activity.

The formal relation between money supply and high-powered money can be stated in the form of equations as under:

The money supply (*M*) consists of deposits of commercial banks (*D*) and currency (*C*) held by the public. Thus the supply of money

$$M=D+C \qquad \ldots(1)$$

High-powered money (*H*) (or monetary base) consists of currency held by the public (*C*) plus required reserves (*RR*) and excess reserves (*ER*) of commercial banks. Thus high-powered money

$$H=C+RR+ER \qquad \ldots(2)$$

The relation between *M* and *H* can be expressed as the ratio of *M* to *H*. So divide equation (1) by (2):

$$\frac{M}{H}=\frac{D+C}{C+RR+ER} \qquad \ldots(3)$$

Divide the numerator and denominator of the right hand side of equation (3) by *D*:

$$\frac{M}{H} = \frac{\frac{D}{D} + \frac{C}{D}}{\frac{C}{D} + \frac{RR}{D} + \frac{ER}{D}}$$

or $$\frac{M}{H} = \frac{1 + \frac{C}{D}}{\frac{C}{D} + \frac{RR}{D} + \frac{ER}{D}} \quad ...(4)$$

By substituting *Cr* for *C*/*D*, *RRr* for *RR*/*D* and *ERr* for *ER*/*D*, equation (4) becomes

$$\frac{M}{H} = \frac{1 + Cr}{Cr + RRr + ERr} \quad ...(5)$$

Thus high-powered money

$$H = \frac{Cr + RRr + ERr}{1 + Cr} \times M \quad ...(6)$$

And money suply

$$M = \frac{1 + Cr}{Cr + RRr + ERr} \times H \quad ..(7)$$

Equation (7) defines money supply in terms of high-powered money. It expresses the money supply in terms of four determinants, *H*, *Cr*, *RRr*, and *ERr*. The equation states that the higher the supply of high powered money, the higher the money supply. Further, the lower the currency ratio (*Cr*), the reserve ration (*RRr*), and the excess reserve ratio (*ERr*), the higher the money supply, and vice versa.

The relation between money supply and high-powered money is illustrated in Figure 1. The horizontal curve *Hs* shows the given supply of high-powered money. The curve *Hd* shows the demand for high-powered money associated with each level of money supply and represents equation (6). The slope of the *Hd* curve is equal to the term (*Cr*+*RRr*+*ERr*)/(1+*Cr*). Given *Cr*, *RRr*, *Err* and the high-powered money Hs, the equilibrium money supply is *OM*. If the money supply is larger

than this, say OM_1, there will be excess demand for high-powered money. On the contrary, a less than *OM* money supply will mean less demand for high-powered money.

If there is an increase in any one of the ratios *Cr* or *RRr* or *ERr*, there would be an increase in the demand for high-powered money. This is shown by the *Hd'* curve in Figure 1 where the increase in the demand for high-powered money leads to decline in the money supply to *OM '*. The quotient of equation (7) is the money multiplier *m*. Thus

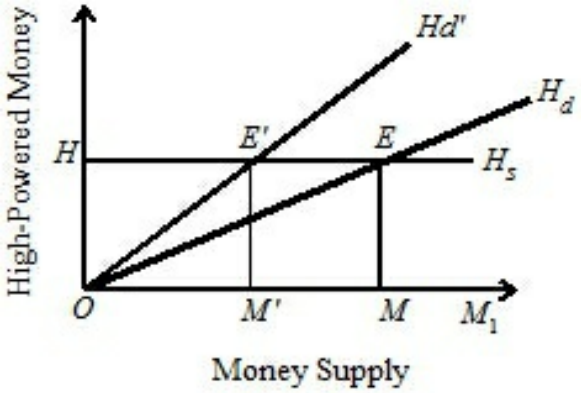


Fig. 1

$$m = \frac{1 + Cr}{CR + RRr + ERr} \qquad \ldots(8)$$

Now the relation between money supply and high-powered money of equation (7) becomes

$$M = mH \qquad \ldots(9)$$

Equation (9) expresses money supply as a function of m and *H*. In other words, the money supply is determined by high powered money (*H*) and the money multiplier (*m*). The size of the money multiplier is determined by the currency ratio (*Cr*) of the public, the required reserve ratio (*RRr*) at the central bank, and the excess reserve ratio (*ERr*) of commercial banks. The lower these ratios are, the larger the money multiplier is. If *m* is fairly stable, the central bank can manipulate the money supply (*M*) by manipulating *H*. The central bank can do so by open market operations. But the stability of *m* depends upon the stability of the currency ratio and the reserve ratios *RRr* and *ERr*. Or, it depends upon off-setting changes in *RRr* and *ERr* ratios. Since these ratios and currency with the public are liable to change, the money multiplier is quite volatile in the short run.

Given the division of high-powered money between currency held by the public, the required reserves at the central bank, and the excess reserves of

commercial banks, the money supply varies inversely with *Cr*, *RRr* and *ERr*. But the supply of money varies directly with changes in the high-powered money. This is shown in Figure 2. An increase in the supply of high-powered money by Δ*H* shifts the *Hs* curve upward to *Hs'*. At *E*, the demand and supply of high-powered money are in equilibrium and money supply is *OM*. With the increase in the supply of high-powered money to *Hs*', the supply of money also increases to *OM* at the new equilibrium point E_1. Further, Figure 2 reveals the operation of the money multiplier. With the increase in the high-powered money Δ*H*, the money supply increases by Δ*M*. An increase in the high-powered money by Re 1 increases by a multiple of Re 1.

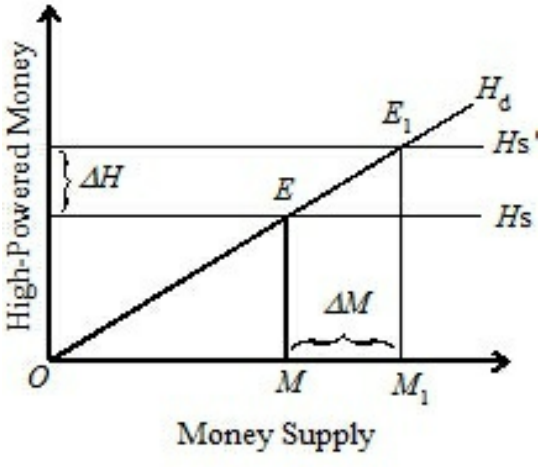


Fig. 2

Some economists do not take into consideration excess reserves in determining high-powered money and consequently the money supply. But the monetarists give more importance to excess reserves. According to them, due to uncertainties prevailing in banking operations as in business, banks always keep excess reserves. The amount of excess reserves depends upon the interaction of two types of costs: the cost of holding excess reserves, and the cost generated by deficiency in excess reserves. The first cost is in terms of the market rate of interest at which excess reserves are maintained. The second cost in in terms of the bank rate which is a sort of penalty to be paid to the central bank for failure to maintain the legal required reserve ratio by the commercial bank. The excess reserve ratio varies inversely with the market rate of interest and direcly with the bank rate. Since the money supply is inversely related to the excess reserve ratio, decline in the excess reserve ratio of banks tends to increase the money supply and vice versa. Thus the money supply is determined by the high-powered money, the currency ratio, the required reserve ratio and the market rate of interest and the bank rate.

The monetary base or high-powered money is directly controllable by the

central bank. It is the ultimate base of the nation's money supply. Of course, the money multiplier times the high-powered money always equals the money supply, *i.e*. $M=mH$. This formulation tells us how much new money will be created by the banking system for a given increase in the high-powered money. The monetary policy of the central bank affects excess reserves and the high-powered money identically. Suppose the central bank makes open market purchases. This raises the high-powered money in the form of excess reserves of banks. An increase in money supply that results from it comes from the banking system which creates new money on the baiss of its newly acquired excess reserves. Thus this concept tells us that the monetary authorities can control the money supply through changing the high-powered money or the money multiplier.

MEASURES OF MONEY SUPPLY IN INDIA

There are four measures of money supply in India which are denoted by M_1, M_2, M_3, and M_4. This classification was introduced by the Reserve Bank of India (RBI) in April 1977. Prior to this till March 1968, RBI published only one measure of the money supply, M or M_1, defined as currency and demand deposits with the public. This was in keeping with the traditional and Keynesian views of the narrow measure of the money supply. From April 1968, RBI also started publishing another measure of the money supply which it called Aggregate Monetary Resources (AMR). This included M_1 *plus* time deposits of banks held by the public. This was a broad measure of money supply which was in line with Friedman's view. But since April 1977, RBI has been publishing data on four measures of the money supply which are discussed as under:

M_1. The first measure of money supply is M_1 which consists of :

(*i*) Currency with the public which includes notes and coins of all denominations in circulation excluding cash on hand with banks:

(*ii*) demand deposits with commercial and cooperative banks, excluding inter-bank deposits; and

(*iii*) 'other deposits' with RBI which include current deposits of foreign central banks, financial institutions and quasi-financial institutions such as IDBI, IFCI, etc, other than of banks, IMF, IBRD, etc. RBI characterises M_1 as *narrow money*.

M_2. The second measure of money supply is M_2 which consists of M_1 *plus* post office savings bank deposits. Since savings bank deposits of commercial and cooperative banks are included in the money supply, it is essential to include post office savings bank deposits. The majority of people in rural and urban India have preference for post office deposits from the safety viewpoint than bank deposits.

M_3. The third measure of money supply in India is M_3 which consists of M_1 *plus* time deposits with commercial and cooperative banks, excluding inter-bank time deposits. RBI calls M_3 as *broad money*.

M_4. The fourth measure of money supply is M_4 which consists of M_3 *plus* total post office deposits comprising time deposits and demand deposits as well. This is the *broadest* measure of money supply.

Of the four inter-related measures of money supply for which RBI publishes data, it is M_3 which is of special significance. It is M_3 which is taken into account in formulating macroeconomic objectives of the economy every year. Since M_1 is narrow money and includes only demand deposits of banks alongwith currency held by the public, it overlooks the importance of time deposits in policy making. That is why, RBI prefers M_3 which includes total deposits of banks and currency with the public in credit budgeting for its credit policy. It is on the estimates of increase in M_3 that the effects of money supply on prices and growth of national income are estimated. In fact M_3 is an empirical measure of money supply in India, as is the practice in developed countries. The Chakravarty Committee also recommended the use of M_3 for monetary targeting without any reason.

MONEY SUPPLY AND LIQUIDITY

Of the four measures of money supply in India, M_1 which consists of currency with the public and demand deposits with commercial and cooperative banks, is the most liquid form of money. Currency consists of notes, rupee coins, two rupee coins, five rupee coins and small coins, and cash on hand with banks, are the most liquid assets. Demand deposits are savings bank accounts and current accounts in banks from which depositors can withdraw cheques by any amount lying in their accounts. Thus like currency, demand deposits are the most liquid and possess the medium of exchange function of money.

A liquid asset is one which is easily spendable, and transferable at face value anywhere and at any time. It can be turned into the generally acceptable medium of exchange quickly without any risk of loss. The phrase 'without risk of loss' refers to the currency unit (Rs, $ or £) and not to real purchasing power. Government bonds, time deposits (also known as savings deposits which are different from savings bank deposits), shares, real estate, etc. are 'frozen' assets which can be sold or exchanged for money on short notice only. They are thus less liquid than money.

M_2 consists of M_1 *plus* post office savings bank deposits. In India, the majority of people in rural and urban areas prefer to keep their cash in post office savings bank deposits from the safety viewpoint because they think that post offices are government owned and mananged. With the nationalisation of 20 commercial banks and opening of their branches in almost all rural areas of the country, the distinction between post office savings bank deposits and commercial savings bank deposits has disappeared. Still the majority of rual people being illiterate, they prefer post offices to banks even by force of habit.

The inclusion of post office savings bank deposits in M_1 is meant to measure the increase in total money supply which affects the evonomy. But post office savings bank deposits are less liquid than currency and demand deposits because they cannot be easily withdrawn. There are no chequing facilities in post offices, except in metropolitan cities and that too in main post offices. The depositors have to undergo a cumbersome process of cash withdrawals in post offices. M_3 includes M_1 plus time deposits (also known as savings deposits in developed countries) with

commercial banks and cooperative banks. This is broad money which stresses the store of value function of money alongwith the medium of exchange function. Time deposits with banks are less liquid than currency and demand deposits because they are held for a fixed time period at a fixed rate of interest. The total money deposited in this account can be withdrawn before the expiry of full period by paying a penal interest rate to the bank. So time deposits do possess liquidity but less than demand deposits.

The fourth measure of money supply is M_4 which includes M_3 *plus* total post office deposits comprising time deposits and savings bank deposits. They tend to increase the money supply in the country manifold. But these total post office deposits are less liquid than total bank deposits for the reasons already given in the case of M_2.

If deposits with non-bank financial institutions such as mutual savings banks, building societies, insurance companies, loan associations and other credit and financial institutions are also included alongwith total post office deposits in M_3 the total money supply would be many times more than what is ordinarily defined as M_1. And if such assets as shares, bonds, government securities, etc. are also included in the supply of money, it would be difficult to measure the money stock in the country.

Taking all such assets vis-a-vis money, they differ in the degree of liquidity. Since *currency* is easily spendable and transferable, and has more stability in value, it possesses the highest degree of liquidity. *Demand deposits* of banks are also as liquid as currency because they are chequing accounts and easily serve as medium of exchange. But demand deposits of post offices do not possess the same degree of liquidity as bank deposits. *Time deposits* of banks, post offices and of other non-bank financial institutions are less liquid than demand deposits because they cannot be easily tranferred to depositors in the form of cash and spent. They serve more as a store of value. So far as *shares* of corporations are concerned, they are also less liquid because they take more time to be sold and transferred. They involve cost in the act of tranferability in the form of brokerage or commission. They cannot be easily converted into cash and spent. Hence they possess less liquidity than demand deposits. *Bonds*

of companies also possess less liquidity because they can be converted into cash after the expiry of the bond maturity period. But they are transferable and earn higher interest return. *Government securities* are issued in the name of initial purchasers and, as such, are non-marketable, because they cannot be sold to someone else. So they are not liquid. On the other hand, money market mutual fund shares, post office savings bonds and national savings certificates possess the advantage of being cashable though they are also non-transferable. They can be returned for repayment of principal plus a fixed amount interest after a short waiting period before the actual maturity date. They are thus as liquid as fixed deposits of banks and post offices.

It is on account these reasons that economists prefer M_1 as the measure of money supply because among all the assets, currency and demand deposits possess the highest degree of liquidity. However, for practical purposes in policy formulation and for empirical studies, governments and researchers use M_3 as the measure of money supply which is less liquid than M_2.

But *how does a change in money supply affect liquidity* ? A change in the money supply affects liquidity by bringing changes or readjustments in the portfolio holdings of the assets of the people. This depends on the effect of money supply on aggregate spending. If people decide to spend the increased money supply in purchasing such assets as shares and debentures, there will be less money available in liquid form with the public. If the stock market is bullish, people may convert assets in their portfolios in buying more shares. On the other hand, if there is uncertainty in the stock market , people may hold the increased money supply in bank deposits or invest it in real estate if they expect property prices to rise. But it is the monetary authority that influences money supply in the economy by following "easy" or "tight" monetary policy. It does so by controlling aggregate spending and thereby influencing business activity, output and employment. But the monetary authority is not always successful in controlling spending by increasing or decreasing the money supply and hence liquidity. This is because the central bank has little control over the velocity of circulation of money, non-bank financial intermediaries, business expectations, time lags in monetary policy, etc. It is, therefore,

very difficult to predict the effects of changes in money supply on liquidity.

DERIVATION OF MONEY MULTIPLIERS

We have explained above the derivation of the money multiplier and its relation with high-powered money. But the total money supply is usually measured not in terms of M but as M_1, M_2, and M_3. Therefore, the derivation of money multipliers of these three measures M_1, M_2, and M_3 of the total money supply and the relation of each with the high powered money are shown below in the form of equations.

First, take M_1. In M_1 measure of money supply are included the demand deposits of commercial banks (D) and currency with the public (C). Thus money supply

$$M_1 = D+C \qquad \ldots(1)$$

In the high-powered money (or monetary base) are included the required reserves of commercial banks (RR) *plus* currency with the public (C). Thus high-powered money

$$H = RR+C \qquad \ldots(2)$$

By dividing equation (1) by equation (2), the ratio of M_1 to H can be expressed as

$$\frac{M_1}{H} = \frac{D+C}{RR+C}$$

Dividing the right hand side of the above equation by D,

$$\frac{M_1}{H} = \frac{\frac{D}{D}+\frac{C}{D}}{\frac{RR}{D}+\frac{C}{D}}$$

or $$\frac{M_1}{H} = \frac{1 + \frac{C}{D}}{\frac{RR}{D} + \frac{C}{D}}$$

By substituting *Cr* for *C*/*D* and *Rr* for *RR*/*D*, the equation becomes

$$\frac{M_1}{H} = \frac{1 + Cr}{Rr + Cr}$$

$$m_1 = \frac{1 + Cr}{Rr + Cr} \times H$$

The higher the value of m_1 multiplier, the lower will be the reserve ratio (*Rr*) and currency-deposit ratio (*Cr*).

M_2 is defined as M_1 *plus* time deposits (*TD*). Thus

$$M_2 = D + C + TD \qquad \text{....(3)}$$

In high-powered money are included required reserves (*RR*), the excess reserves of commercial banks (*ER*) *plus* currency held by the public (*C*). So

$$H = RR + Er + C \qquad \text{...(4)}$$

In order to find out the ratio of M_2 to *H*, divide equation (3) by equation (4),

$$\frac{M_2}{H} = \frac{D + C + TD}{RR + ER + C}$$

Dividing the right hand side of the above equation by *D*,

$$\frac{M_2}{H} = \frac{\frac{D}{D} + \frac{C}{D} + \frac{TD}{D}}{\frac{RR}{D} + \frac{ER}{D} + \frac{C}{D}}$$

By substituting *Cr* for *C*/*D*, *Td* for *TD*/*D*, *RRr* for *RR*/*D* and *Er* for *ER*/*D*,

$$\frac{M_2}{H} = \frac{1 + Cr + Td}{Rr + Er + Cr}$$

Or $$m_2 = \frac{1 + Cr + Td}{Rr + Er + Cr} \times H$$

The value of m_2 multiplier is higher than that of m_1 multiplier because it leads to greater increase in the monetary base (*H*). The higher the value of m_2 multiplier, the lower will be the rates of *Cr*, *Rr* and *Er*.

M_3 includes M_2 *plus* deposits of non-bank financial institutions (*Dn*) and a portion of deposits of these institutions which remains with banks (*Rp*). Thus the deposits of these institutions are taken as *RpDn* which are related to the reserve ratio of commercial banks (*Rr*). Thus

$$M_3 = D + RpDn + Dn + Cr\ (D + Dn) \qquad ...(5)$$

And the monetary base will be determined as :

$$H = Rr\ (D + RpDn) + Cr(D + Dn) \qquad ...(6)$$

$Rr = RR/D$ or $R = Rr\ (D)$

$Cr = C/D$ or $C = Cr\ (D)$

To find out the ratio of M_3 and *H*, divide equation (5) by equation (6),

$$\frac{M_3}{H} = \frac{D + RpDn + Dn + Cr(D + Dn)}{Rr(D + RpDn) + Cr(D + Dn)}$$

Or $$\frac{M_3}{H} = \frac{D + Dn(Rp + 1) + Cr(D + Dn)}{Rr(D + RpDn) + Cr(D + Dn)}$$

Dividing the right hand side of the above equation by *D*.

$$\frac{M_3}{H} = \frac{1 + \frac{Dn}{D}(Rp + 1) + Cr\left(1 + \frac{Dn}{D}\right)}{Rr\left(1 + Rp\frac{Dn}{D}\right) + Cr\left(1 + \frac{Dn}{D}\right)}$$

$$= \frac{1 + \frac{Dn}{D}Rp + \frac{Dn}{D} + Cr + Cr \times \frac{Dn}{D}}{Rr + Rr.Rp.\frac{Dn}{D} + Cr \times \frac{Dn}{D}}$$

$$= \frac{1 + Cr + \frac{Dn}{D}Rp + \frac{Dn}{D} + Cr \times \frac{Dn}{D}}{Rr + Cr + Rr.Rp.\frac{Dn}{D} + Cr \times \frac{Dn}{D}}$$

$$= \frac{1 + Cr + \frac{Dn}{D}(Rp + 1 + Cr)}{Rr + Cr + \frac{Dn}{D}(Rp + 1 + Cr)}$$

By substituting *d* for *Dn*/*D*, the equation becomes

$$\frac{M_3}{H} = \frac{1 + Cr + (1 + Rp + Cr)d}{Rr + Cr + (Rr.Rp + Cr)d}$$

Or $$m_3 = \frac{1 + Cr + (1 + Rp + Cr)d}{Rr + Cr + (Rr.Rp + Cr)d} \times \mathrm{H}$$

where m_3 is the money multiplier whose valve is higher than m_2 multiplier.

EXERCISES

1. Explain the various components of the money supply.
2. Discuss the determinants of the money supply. Should time deposits be included under the supply of money ?
3. Discuss the various measures of the money supply adopted by the Reserve Bank of India. How do they differ from those adopted by

developed countries?

4. Of the various measures of money supply which possess more liquidity and why?
5. Explain the determinants of high-powered money.
6. Discuss the relation between money supply and high-powered money.
7. Write a note on the money multiplier.

CHAPTER 29

CREDIT CREATION BY COMMERCIAL BANKS

DO BANKS CREATE CREDIT?

The creation of credit or deposits is one of the most important functions of commercial banks. Like other corporations, banks aim at earning profits. For this purpose, they accept cash in demand deposits and advance loans on credit to customers. When a bank advances a loan, it does not pay the amount in cash. But it opens a current account in his name and allows him to withdraw the required sum by cheques. In this way, the bank creates credit or deposits.

Demand deposits arise in two ways: *one*, when customers deposit currency with commercial banks, and *two*, when banks advance loans, discount bills, provide overdraft facilities, and make investments through bonds and securities. The first type of demand deposits are called "primary deposits". Banks play a passive role in opening them. The second type of demand deposits are called "derivative deposits". Banks actively create such deposits.

Do banks really create credit or deposits ?

There have been two views on this subject: one held by certain economists like Hartley Withers, and the other held by practical bankers like Walter Leaf.

---

1. Water Leaf, *Banking*, 1928, pp. 101-4.

According to Withers, banks can create credit by opening a deposit, every time they advance a loan. This is because every time a loan is sanctioned, payment is made through cheques by the customers. All such payments are adjusted through the clearing house. So long as a loan is due, a deposit of that amount remains outstanding in the books of the bank. Thus every loan creates a deposit. But this is an exaggerated and extreme view.

Dr Leaf[1] and practical bankers do not agree with this view. They go to the opposite extreme. They hold that banks cannot create money out of thin air. They can lend only what they have in cash. Therefore, they cannot and do not create money.

This view is also wrong becuase it is based on arguments relating to a single bank. As pointed out by Prof Samuelson, "The banking system as a whole can do what each small bank cannot do: it can expand its loans and investments many times the new reserves of cash created for it, even though each small bank is lending out only a fraction of its deposits."[2]

In fact, a bank is not a cloak room where one can keep currency notes and claim those very notes when one desires. Banks know by experience that all depositors do not withdraw their money simultaneously. Some withdraw while others deposit on the same day. So by keeping a small cash in reserve for day-to-day transactions, the bank is able to advance loans on the basis of excess reserves. When the bank advances a loan it opens an account in the name of the customer. The bank knows by experience that the customer will withdraw money by cheques which will be deposited by his creditors in this bank or some other bank, where they have their accounts. Settlements of all such cheques are made in the clearing house. The same procedure is followed in other banks. The banks are able to create credit or deposits by keeping a small cash in reserves and lending the remaining amount.

In granting a loan, a bank actively creates a claim against itself in favour of the borrower. "The claims the bank takes from its customers, in exchange for the deposits entered in the books, are the bank's assets. The standard assets of a commercial bank are overdrafts and loans, bills discounted, investments and cash."[3]

The bank provides overdraft facility to a customer on the basis of some security. It enters the amount of the overdraft in the existing account of the customer and allows him to draw cheques for the overdraft amount agreed upon. It thus creates a deposit.

When a bank discounts a bill of exchange, it in fact, buys the bill from the customer for a short period of 90 days or less. The amount of the bill is credited in the account of the customer who withdraws it through a cheque. Or, it pays the sum through a cheque on itself. In both cases, the bank creates a deposit equal to the amount of the bill of exchange less the discount charges.

A commercial bank also creates a deposit by making investments by buying government bonds and securities. The bank pays for the bond through a cheque on itself to the central bank. If it buys a bond from the stock exchange, it credits the amount in the account of the seller, if he happens to be its customer. Otherwise, it pays a cheque on itself which is deposited in some other bank. In any case, a deposit is created either in this bank or some other bank. In all such cases, liabilities and assets in the banking system on the whole are increased. Thus loans by banks create deposits. It is in this sense that credit is created by commercial banks.

---

2. Paul A. Samuelson, *Economics*, 10th edn., 1976, p. 301.

3. R.S. Sayers, *Modern Banking*, 7th edn,. 1967, pp. 7-8.

THE PROCESS OF CREDIT CREATION

Let us explain tha actual process of credit creation. We have seen above that the ability of banks to create credit depends on the fact that banks need only a small percentage of cash to deposits. If banks kept 100 per cent cash against deposits, there would be no credit creation. Modern banks do not keep 100 per cent cash reserves. They are legally required to keep a fixed percentage of their deposits in cash, say 10, 15 or 20 per cent. They lend and/or invest the remaining amount which is called *excess reserves*. A bank can lend equal to its excess reserves. But the entire banking system can lend and create credit (or deposits) upto a multiple of

its original excess reserves. The deposit multiplier depends upon the required reserve which is the basis of credit creation. Symbolically, the required reserve ratio:

$$RRr = \frac{RR}{D}$$
$$\text{or } RR = RRr \times D$$

where *RR* are the required cash reserves with banks, *RRr* is the required reserve ratio and *D* is the demand deposits of banks. To show that *D* depends on *RR* and *RRr*, divide both sides of the above equation by *RRr*:

$$\frac{RR}{RRr} = \frac{RRr \times D}{RRr}$$
$$\text{or } \frac{RR}{RRr} = D$$
$$\text{or } \frac{1}{RRr} = \frac{D}{RR}$$
$$\text{or } D = \frac{1}{RRr} \times RR$$

where 1/*RRr*, the reciprocal of the percentage reserve ratio, is called the *deposit (or credit) expansion multiplier*. It determines the limits of the deposit expansion of a bank. The maximum amount of demand deposits which the banking system can support with any given amount of *RR* is by applying the multiplier to *RR*. Taking the initial change in the volume of deposits (ΔD) and in cash reserves (ΔRR), it follows from any given percentage of *RRr* that

$$\Delta D = RR \times \frac{1}{RRr}$$

To understand it, suppose the *RRr* for the banks is fixed at 10 per cent and the initial change in cash reserves is Rs 1000. By applying the above formula, the maximum increase in demand deposits will be

$$\Delta D = 1000 \times \frac{1}{0.10} = \text{ Rs } 10000.$$

This is the extent to which the banking system can create credit. The above equation can also be expressed as follows:

$\Delta D = RR[1+(1-RRr) + (1-RRr)^2+...+(1-RRr)^n]$

The sum of the geometric progression within brackets gives:

$$\frac{1}{1-(1-RRr)} = \frac{1}{RRr}$$

$$\Delta D = \Delta RR \times \frac{1}{RRr}$$

The deposit expansion multiplier rests on the assumptions that banks lend out all their excess reserves and *RRr* remains constant.

To explain the process of credit creation, we make the following assumptions:

1. There are many banks, say A, B, C, etc. in the banking system.
2. Each bank has to keep 10 per cent of its deposits in reserves. In other words, 10 per cent is the required ratio fixed by law.
3. The first bank has Rs. 1000 as deposits.
4. The loan amount drawn by the customer of one bank is deposited in full in the second bank, and that of the second bank into the third bank, and so on.
5. Each bank starts with the initial deposit which is deposited by the debtor of the other bank.

Given these assumptions suppose that Bank A receives a cash deposits of Rs. 1000 to begin with. This is the cash in hand with the bank which is its asset and this amount is also the liability of the bank by way of deposits it holds. Given the reserve ratio of 10 per cent, the bank keeps Rs. 100 in reserves and lends Rs 900 to one of its customers who, in turn, gives a cheque to some person from whom he borrows or buys something. The

net changes in Bank A' is balance sheet are +Rs 100 in reserves and +Rs 900 in loans on the assets side and Rs 1000 in demand deposits on the liabilities side as shown in Table I. Before these changes Bank A had zero excess reserves.

Table I. Balance Sheet of Bank A

Assets		**Liabilities**	
Reserves	Rs 1000	Deposits	Rs 1000
net changes		*net changes*	
Reserves	Rs 1000	Deposits	Rs 1000
Loans	Rs 900		

This loan of Rs 900 is deposited by the customer in Bank *B* whose balance sheet is shown in Table II Bank *B* starts with a deposit of Rs 900, keeps 10 per cent of it or Rs 90 as cash in reserve. Bank *B* has Rs 810 as excess reserves which it lends thereby creating new deposits.

Table II Balance Sheet of Bank B

Assets		**Liabilities**	
Reserves	Rs 900	Deposits	Rs 900
net changes		*net changes*	
Reserves	Rs 90	Deposits	Rs 900
Loans	Rs 810		

This loan of Rs 810 is deposited by the customer of Bank *B* into Bank *C*.

The balance sheet of Bank *C* is shown in Table III Bank *C* keeps Rs 81 or 10 per cent of Rs 810 in cash reserves and lends Rs 729.

Table III Balance Sheet of Bank C

Assets		**Liabilities**	
Reserves	Rs 810	Deposits	Rs 810
net changes		*net changes*	
Reserves	Rs 81	Deposits	Rs 810
Loans	Rs 729		

This process goes on to other banks. Each bank in the sequence gets excess reserves, lends and creates new demand deposits equal to 90% of the preceding bank's. In this way, new deposits are created to the tune of Rs 10000 in the banking system, as shown in Table IV.

Table IV Multiple Credit Creation

Bank	Required Reserves	New Loans	New Deposits
A	Rs 100	Rs 900	Rs 1000
B	Rs 90	Rs 810	Rs 900
C	Rs 81	Rs 729	Rs 810
All other Banks	Rs 729	Rs 6561	Rs 7290
Total for the Banking System	Rs 1000	Rs 9000	Rs 10000

The multiple credit creation shown in the last column of the above table can also be worked out algebraically as:

Rs 1000 [1+(9/10)+(9/10)2+(9/10)3+...+(9/10)"]

=Rs 1000 (1/1-9/10)= Rs 1000 (1/1/10)= Rs 1000×10 = Rs 10000.

LIMITATIONS ON THE POWER OF BANKS TO CREATE CREDIT

We have seen above how the banking system as a whole can create credit. But it does not mean that banks have unlimited powers to create credit. In fact, they have to function under certain restrictions. The following are the limitations on the power of commercial banks to create credit.

1. Amount of Cash. The credit creation power of banks depends upon the amount of cash they possess. The larger the cash, the larger the amount of credit that can be created by banks. The amount of cash that a bank has in its vaults cannot be determined by it. It depends upon the primary deposits with the bank. The bank's power of creating credit is thus limited by the cash it possesses.

2. Proper securities. An important factor that limits the power of a bank to create credit is the availability of adequate securities. A bank advances loans to its customers on the basis of a security, or a bill, or a share, or a stock or a building, or some other type of asset. It turns ill-liquid form of wealth into liquid wealth and thus creates credit. If proper securities are not available with the public, a bank cannot create credit. As pointed out by Crowther, "Thus the bank does not create money out of thin air, it transmutes other forms of wealth into money."

3. Banking habits of the People. The banking habits of the people also govern the power of credit creation on the part of banks. If people are not in the habit of using cheques, the grant of loans will lead to the withdrawal of cash from the credit creation stream of the banking system. This reduces the power of banks to create credit to the desired level.

4. Minimum Legal Reserve Ratio. The minimum legal reserve ratio of cash to deposits fixed by the central bank is an important factor which determines the power of banks of create credit. The higher this ratio (RRr), the lower the power of banks to create credit; and the lower the

ratio, the higher the power of banks to create credit.

5. Excess Reserves. The process of credit creation is based on the assumption that banks stick to the required reserve ratio fixed by the central bank. If banks keep more cash in reserves than the legal reserve requirements, their power to create credit is limited to that extent. If Bank A of our example keeps 25 per cent of Rs 1000 instead of 20 per cent, it will lend Rs 750 instead of Rs 800. Consequently, the amount of credit creation will be reduced even if the other banks in the system stick to the legal reserve ratio of 20 per cent.

6. Leakages. If there are leakages in the credit creation stream of the banking system, credit expansion will not reach the required level, given the legal reserve ratio. It is possible that some persons who receive cheques do not deposit them in their bank accounts, but withdraw the money in cash for spending or for hoarding at home. The extent to which the amount of cash is withdrawn from the chain of credit expansion, the power of the banking system to create credit is limited.

7. Cheque Clearances. The process of credit expansion is based on the assumption that cheques drawn by commercial banks are cleared immediately and reserves of commercial banks expand and contract uniformly by cheque transactions. But it is not possible for banks to receive and draw cheques of exactly equal amount. Often some banks have their reserves increased and otheres reduced through cheque clearances. This expands and contracts credit creation on the part of banks. Accordingly, the credit creation stream is disturbed.

8. Behaviour of other Banks. The power of credit creation is further limited by the behaviour of other banks. If some of the banks do not advance loans to the extent required of the banking system, the chain of credit expansion will be broken. Consequently, the banking system will not be "loaned up".

9. Economic Climate. Banks cannot continue to create credit limitlessly. Their power to create credit depends upon the economic climate in the country. If there are boom times, there is optimism. Investment opportunities increase and businessmen take more loans from banks. So

credit expands. But in depressed times when the business activity is at a low level, banks cannot force the business community to take loans from them. Thus the economic climate in a country determines the power of banks to create credit.

10. Credit Control Policy of the Central Bank. The power of commercial banks to create credit is also limited by the credit control policy of the central bank. The central bank influences the amount of cash reserves with banks by open market operations, discount rate policy and varying margin requirements. Accordingly, it affects the credit expansion or contraction by commercial banks.

We may conclude that commercial banks do not possess unlimited powers to create credit.

EXERCISES

1. Explain the statement that "loans create deposits."
2. Explain how "loans make deposits". What are the limitations to such credit creation by banks ?
3. Describe the process of credit creation by commercial banks.
4. How do banks create credit ? What are the limitations on the power of banks to create credit ?

CHAPTER 30

CENTRAL BANKING: FUNCTIONS AND CREDIT CONTRAOL

INTRODUCTION

The central bank is the apex bank in a country. It is called by different names in different countries. It is the Reserve Bank of India in India, the Bank of England in England, the Federal Reserve System in America, the Bank of France in France, the Riksbank in Sweden, etc.

DEFINITION OF A CENTRAL BANK

A central bank has been defined in terms of its functions. According to Vera Smith, "The primary definition of central banking is a banking system in which a single bank has either complete control or a residuary monopoly of note issue." W.A. Shaw defines a central bank as a bank which controls credit. To Hawtrey, a central bank is that which is the lender of the last resort. According to A.C.L. Day, a central bank is "to help control and stabilise the monetary and banking system." According to Sayers, the central bank "is the organ of government that undertakes the major financial operations of the government and by its conduct of these operations and by other means, influences the behaviour of financial institutions so as to support the economic policy of the government." Sayers refers only to the nature of the central bank as the government's bank. All these definitions are narrow because they refer only to one particular function of a central bank.

On the other hand, Samuelson's definition is wide. According to him, a central bank "is a bank of bankers. Its duty is to control the monetary base.... and through control of this 'high-powered money' to control the community's supply of money." But the broadest definition has been given by De Kock in his book *Central Banking*. In his words, a central bank is "a bank which constitutes the apex of the monetary and banking structure of its country and which performs as best as it can in the national economic interest, the following functions: (i) The regulation of currency in accordance with the requirements of business and the general public for which purpose it is granted either the sole right of note issue or at least a partial monopoly thereof. (ii) The performance of general banking and agency for the state. (iii) The custody of the cash reserves of the commercial banks. (iv) The custody and management of the nation's reserves of international currency. (v) The granting of accommodation in the form of re-discounts and collateral advances to commercial banks, bill brokers and dealers, or other financial institutions and the general acceptance of the responsibility of lender of the last resort. (vi) The settlement of clearance balances between the banks. (vii) The control of credit in accordance with the needs of business and with a view to carrying out the broad monetary policy adopted by the state." De Kock's definition is too long to be called a definition. For, a definition must be brief.

FUNCTIONS OF A CENTRAL BANK

A central bank performs the following functions, as given by De Kock and accepted by the majority of economists.

1. Regulator of Currency

The central bank is the bank of issue. It has the monopoly of note issue. Notes issued by it circulate as legal tender money. It has its issue department which issues notes and coins to commercial banks. Coins are manufactured in the government mint but they are put into circulation through the central bank.

Central banks have been following different methods of note issue in

different countries. The central bank is required by law to keep a certain amount of gold and foreign securities against the issue of notes. In some countries, the amount of gold and foreign securities bears a fixed proportion, between 25 to 40 per cent of the total notes issued. In other countries, a minimum fixed amount of gold and foreign currencies is required to be kept against note issue by the central bank. This system is operative in India whereby the Reserve Bank of India is required to keep Rs 115 crores in gold and Rs 85 crores in foreign securities. There is no limit to the issue of notes after keeping this minimum amount of Rs 200 crores in gold and foreign securities.

The monopoly of issuing notes vested in the central bank ensures uniformity in the notes issued which helps in facilitating exchange and trade within the country. It brings stability in the monetary system and creates confidence among the public. The central bank can restrict or expand the supply of cash according to the requirements of the economy. Thus it provides elasticity to the monetary system. By having a monopoly of note issue, the central bank also controls the banking system by being the ultimate source of cash. Last but not the least, by entrusting the monopoly of note issue to the central bank, the government is able to earn profits from printing notes whose cost is very low as compared with their face value.

2. Banker, Fiscal Agent and Adviser to the Governement

Central banks everywhere act as bankers, fiscal agents and advisers to their respective governments. As *banker* to the government, the central bank keeps the deposits of the central and state governments and makes payments on behalf of governments. But it does not pay interest on governments depostis. It buys and sells foreign currencies on behalf of the government., It keeps the stock of gold of the government. Thus it is the custodian of government money and wealth. As a *fiscal agent*, the central bank makes short-term loans to the government for a period not exceeding 90 days. It floats loans, pays interest on them, and finally repays them on behalf of the government. Thus it manages the entire *public debt*. The central bank also *advises* the government on such economic and money matters as controlling inflation or deflation, devaluetion or revaluation of

the currency, deflicit financing, balance of payments, etc. As pointed out by De Kock, "Central banks everywhere operate as bankers to the state not only because it may be more convenient and economical to the state, but also because of the intimate connection between public finance and monetary affairs."

3. Custodian of Cash Reserves of Commercial Banks

Commercial banks are required by law to keep reserves equal to a certain percentage of both time and demand desposits liabilities with the central bank. It is on the basis of these reserves that the central bank transfers funds from one bank to another to facilitate the clearing of cheques. Thus the central bank acts as the custodian of the cash reserves of commercial banks and helps in facilitating their transactions. There are many advantages of keeping the cash reserves of the commercial banks with the central bank. *Firstly,* the centralisation of cash reserves in the central bank is a source of great strength to the banking system of a country. *Secondly,* centralised cash reserves can serve as the basis of a large and more elastic credit structure than if the same amount were scattered among the individual banks. *Thirdly,* centralised cash reserves can be utilised fully and most effectively during periods of seasonal strains and in financial crises or emergencies. *Fourthly,* by varying these cash reserves the central bank can control the credit creation by commercial banks. *Lastly,* the central bank can provide additional funds on a temporary and short term basis to commercial banks to overcome their financial difficulties.

4. Custody and Management of Foreign Exchange Reserves

The central bank keeps and manages the foreign exchange reserves of the country. It is an official reservoir of gold and foreign currencies. It sells gold at fixed prices to the monetary authorities of other countries. It also buys and sells foreign currencies at international prices. Further, it fixes the exchange rates of the domestic currency in terms of foreign currencies. It holds these rates within narrow limits in keeping with its obligations as a member of the International Monetary Fund and tries to bring stability in foreign exchange rates. Further, it manages *exchange control operations* by supplying foreign currencies to importers and persons visiting foreign countries on business, studies, etc. in keeping with the rules laid down by

the government.

5. Lender of the Last Resort

De Kock regards this function as a *sine qua non* of central banking. By granting accommodation in the form of re-discounts and collateral advances to commercial banks, bill brokers and dealers, or other financial institutions, the central bank acts as the lender of the last resort. The central bank lends to such institutions in order to help them in times of stress so as to save the financial structure of the country from collapse. It acts as lender of the last resort through discount house on the basis of treasury bills, government securities and bonds at the "front door". The other method is to give temporary accommodation to the commercial banks or discount houses directly through the "back door". The difference between the two methods is that lending at the front door is at the bank rate and in the second case at the market rate. Thus the central bank as lender of the last resort is a big source of cash and also influences prices and market rates.

6. Clearing House for Transfer and Settlement

As bankers' bank, the central bank acts as a clearing house for transfer and settlement of mutual claims of commercial banks. Since the central bank holds reserves of commercial banks, it transfers funds from one bank to other banks to facilitate clearing of cheques. This is done by making transfer entries in their accounts on the principle of book-keeping. To transfer and settle claims of one bank upon others, the central bank operates a seperate department in big cities and trade centres. This department is known as the "clearing house" and it renders the service free to commercial banks.

When the central bank acts as a clearing agency, it is time-saving and convenient for the commercial banks to settle their claims at one place. It also economises the use of money. "It is not only a means of economising cash and capital but is also a means of testing at any time the degree of liquidity which the community is maintaining."[1]

7.Controller of Credit

The most important function of the central bank is to control the credit creation power of commercial bank in order to control inflationary and deflationary pressures within the economy. For this purpose, it adopts quantitative methods and qualitative methods. *Quantitative methods* of control aim at controlling the cost and quantity of credit by adopting bank rate policy, open market operations, and by variations in reserve ratios of commercial banks. *Qualitative methods* control the use and direction of credit. These involve selective credit controls and direct action. By adopting such methods, the central bank tries to influence and control credit creation by commercial banks in order to stabilise economic activity in the country.

8. Other Functions

Besides the above noted functions, the central banks in a number of developing countries have been entrusted with the responsibility of developing a strong banking system to meet the expanding requirements of agriculture, industry, trade and commerce. Accordingly, the central banks possess some additional powers of supervision and control over the commercial banks. They are the issuing of licences; the regulation of branch expansion; to see that every bank maintains the minimum paid up capital and reserves as provided by law; inspecting or auditing the accounts of banks; to approve the appointment of chairmen and directors of such banks in accordance with the rules and qualifications; to control and recommend merger of weak banks in order to avoid their failures and to protect the interest of depositors; to recommend nationalisation of certain banks to the government in public interest; to publish periodical reports relating to different aspects of monetary and economic policies for the benefit of banks and the public; to engage in research; and to train banking personnel, etc.

1.H.P. Willis, *Theory and Practice of Central Banking*, p. 313.

CENTRAL BANK AS THE CONTROLLER OF CREDIT

Objectives of Credit Control

Credit control is the means to control the lending policy of commercial banks by the central bank. The central bank controls credit to achieve the following objectives:

1. To Stabilise the Internal Price Level. One of the objectives of controlling credit is to stabilise the price level in the country. Frequent changes in prices adversely affect the economy. Inflationary or deflationary trends need to be prevented. This can be achieved by adopting a judicious policy of credit control.

2. To Stabilise the Rate of Foreign Exchange. With the change in the internal price level, exports and imports of the country are affected. When prices fall, exports increase and imports decline. Consequently, the demand for domestic currency increases in the foreign market and its exchange rate rises. On the contrary, a rise in domestic prices leads to a decline in exports and an increase in imports. As a result, the demand for foreign currency increases and that of domestic currency falls, thereby lowering the exchange rate of the domestic currency. Since it is the volume of credit money that affects prices, the central bank can stabilise the rate of foreign exchange by controlling bank credit.

3. To Protect the Outflow of Gold. The central bank holds the gold reserves of the country in its vaults. Expansion of bank credit leads to rise in prices which reduce exports and increase imports, thereby creating an unfavourable balance of payments. This necessitates the export of gold to other countries. The central bank has to control credit in order to prevent such outflows of gold to other countries.

4. To Control Business Cycles. Business cycles are a common phenomenon of capitalist countries which lead to periodic fluctuations in production, employment and prices. They are characterised by alternating periods of prosperity and depression. During prosperity, there is large expansion in the volume of credit, and production, employment and prices rise. During depression, credit contracts, and production, employment and prices fall. The central bank can counteract such cyclical fluctuations through contraction of bank credit during boom periods, and expansion of bank credit during depression.

5. To Meet Business Needs. According to Burgess, one of the important objectives of credit control is the "adjustment of the volume of credit to the volume of business." Credit is needed to meet the requirements of trade and industry. As business expands, larger quantity of credit is needed, and when business contracts less credit is needed. Therefore, it is the central bank which can meet the requirements of business by controlling credit.

6. To Have Growth with Stability. In recent years, the principal objective of credit control is to have growth with stability. The other objectives, such as price stability, foreign exchange rate stability, etc., are regarded as secondary. The aim of credit control is to help in achieving full employment and accelerated growth with stability in the economy without inflationary pressures and balance of payments deficits.

Methods of Credit Control

The central bank adopts two types of methods of credit control. They are the quantitative and qualitative methods. *Quantitative* methods aim at controlling the cost and quantity of credit by adopting such techniques as variations in the bank rate, open market operations, and variations in the reserve ratios of commercial banks. On the other hand, *qualitative* methods control the use and direction of credit. These involve selective credit controls and direct action.

Quantitative Methods.

The quantitative methods of credit control are discussed below.

1. Bank Rate or Discount Rate Policy

The bank rate or the discount rate is the rate fixed by the central bank at which it rediscounts first class bills of exchange and government securities held by the commercial banks. The bank rate is the interest rate charged by the central bank at which it provides rediscount to banks through the *discount window*. The central bank controls credit by making variations in the bank rate. If the need of the economy is to expand credit, the central bank lowers the bank rate. Borrowing from the central bank

becomes cheap and easy. So the commercial banks will borrow more. They will, in turn, advance loans to customers at a lower rate. The market rate of interest will be reduced. This encourages business activity, and expansion of credit follows which encourages the rise in prices. The opposite happens when credit is to be contracted in the economy.

Limitations of Bank Rate Policy

The efficacy of the bank rate policy as an instrument of controlling credit is limited by the following factors:

1. Market Rates do not change with Bank Rate. The success of the bank rate policy depends upon the extent to which other market rates of interest change along with the bank rate. The theory of bank rate policy pre-supposes that other rates of interest prevailing in the money market change in the direction of the change in the bank rate. If this condition is not satisfied, the bank rate policy will be totally ineffective as an instrument of credit control.

2. Wages, Costs and Prices not Elastic. The success of the bank rate policy requires elasticity not only in interest rates but also in wages, costs and prices. It implies that when suppose the bank rate is raised, wages, costs and prices should automatically adjust themselves to a lower level. But this was possible only under gold standard. Now-a-days the emergence of strong trade unions have made wages rigid during deflationary trends. And they also lag behind when there are inflationary tendencies because it takes time for unions to get a wage rise from employers. So the bank rate policy cannot be a success in a rigid society.

3. Banks do not approach Central Bank. The effectiveness of the bank rate policy as a tool of credit control is also limited by the behaviour of the commercial banks. It is only if the commercial banks approach the central bank for rediscounting facilities that this policy can be a success. But the banks keep with them large amounts of liquid assets and do not find it necessary to approach the central bank for financial help.

4. Bills of Exchange not Used. As a corollary to the above, the effectiveness of the bank rate policy depends on the existence of eligible

bills of exchange. In recent years, the bill of exchange as an instrument of financing commerce and trade has fallen into disuse. Businessmen and banks prefer cash credit and overdrafts. This makes the bank rate policy less effective for controlling credit in the country.

5. Pessimism or Optimism. The efficacy of the bank rate policy also depends on waves of pessimism or optimism among businessmen. If the bank rate is raised, they will continue to borrow even at a higher rate of interest if there are boom conditions in the economy, and prices are expected to rise further. On the other hand, a reduction in the bank rate will not induce them to borrow during periods of falling prices. Thus businessmen are not very sensitive to changes in interest rates and they are influenced more by business expectations.

6. Power to Control Deflation Limited. Another limitation of the bank rate policy is that the power of a central bank to force a reduction in the market rates of interest is limited. For instance, a lowering of bank rate below 3 per cent will not lead to a decline in the market rates of interest below 3 per cent. So the bank rate policy is ineffective in controlling deflation. It may, however, control inflationary tendencies by forcing an increase in the market rates of interest.

7. Level of Bank Rate in relation to Market Rate. The efficacy of the discount rate policy as an instrument of credit control depends upon its level in relation to the market rate. If in a boom the bank rate is not raised to such an extent as to make borrowing costly from the central bank, and it is not lowered during a recession so as to make borrowing cheaper from it, it would have a destabilising effect on economic activity.

8. Non-Discriminatory. The bank rate policy is non-discriminatory because it does not distinguish between productive and unproductive activities in the country.

9. Not Successful in Controlling BOP Disequilibrium. The bank rate policy is not effective in controlling balance of payments disequilibrium in a country because it requires the removal of all restrictions on foreign exchange and movements of international capital.

Conclusion. The above points have led the majority of economists to conclude that the power to alter the bank rate is an extremely weak weapon of monetary management. Friedman even went to the extent of proposing outright elimination of the "discount window" itself.

2. Open Market Operations

Open market operations are another method of quantitative credit control used by a central bank. This method refers to the sale and purchase of securities, bills and bonds of government as well as private financial institutions by the central bank. But in its narrow sense, it simply means dealing only in government securities and bonds.

There are two principal motives of open market operations. *One*, to influence the reserves of commercial banks in order to control their power of credit creation. *Two*, to affect the market rates of interest so as to control the commercial bank credit.

Its method of operation is as follows. Suppose the central bank of a country wants to control expansion of credit by the commercial banks for the purpose of controlling inflationary pressures within the economy. It sells government securities in the money market amounting to, say, Rs 10 crores. The buyers of securities give the central bank cheques for this amount drawn against the commercial banks in which they have their accounts. The central bank reduces this amount in their accounts with it. This applies equally if the commmercial banks have also purchased securities from the central bank. The sale of securities in the open market has thus reduced their cash holdings with the central bank. This tends to reduce the actual cash ratio of the commercial banks by Rs. 10 crores. So the banks are forced to curtail their lending.

Suppose that initially the commercial banks have assets worth Rs 1000 crores and the cash-deposit ratio is 10. This means that Rs 1000 crores are divided between Rs 100 crores cash and Rs 900 crores as deposits or loans. As a result of the sale of securities worth Rs 10 crores by the central bank, the cash is reduced by Rs 100 crores. So the total cash with banks remains Rs 900 crores and loans are also reduced by the same percentage, that is, to Rs 810 crores. This reduction in the cash holding of the

commercial banks causes a decrease in the supply of bank money, as shown in Figure 1. In this figure, *S* is the supply curve of bank money which shifts to the left as S_1showing a decrease in the supply of bank money from *B* to *A*, given the level of interest rate *r*.

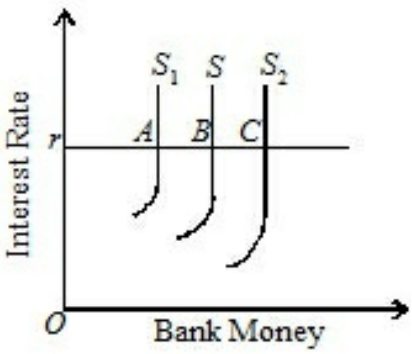


Fig. 1

On the other hand, when the central bank aims at an expansionary policy during a recessionary period, it purchases government securities from the commercial banks and institutions dealing with such securities. The central bank pays the sellers its cheques drawn against itself which are deposited into their accounts with the commercial banks. The reserves of the latter increase with the central bank which are just like cash. As a result, the supply curve of bank money shifts to the right from *S* to S_2showing an increase in the supply of bank money from *B* to *C*, as shown in Figure 1. The banks will now lend more at the given rate of interest, *r*.

Another aspect of the open market policy is that when the supply of money changes as a result of open market operations, the market rates of interest also change. A decrease in the supply of bank money through the sale of securities will have the effect of raising the market interest rates. On the other hand, an increase in the supply of bank money through the purchase of securities will reduce the market interest rates. Thus open market operations have a direct influence on the market rates of interest also.

Limitations of Open Market Operations

The effectiveness of open market operations as a method of credit control is dependent upon the existence of a number of conditions, which limit the full working of this policy.

1. Lack of Securities Market. The first condition is the existence of a large and well-organised security market. This condition is very essential for open market operations because without a well developed security

market the central bank will not be able to buy and sell securities on a large scale, and thereby influence the reserves of the commercial banks. Further, the central bank must have enough saleable securities with it.

2. Cash Reserve Ratio not Stable. The success of open market operations also requires the maintenance of a stable cash reserve ratio by the commercial bank. It implies that when the central bank sells or buys securities, the reserves of the commercial banks decrease or increase accordingly to maintain the fixed ratio. But usually the banks do not stick to the legal minimum reserve ratio and keep a higher ratio than this. This makes open market operations less effective in controlling the volume of credit.

3. Penal Bank Rate. According to Prof. Aschheim, one of the necessary conditions for the success of open market operations is a penal bank rate. If there is no penal discount rate fixed by the central bank, the commercial banks can increase their borrowings from it when the demand for credit is strong on the part of the latter. In this situation, the scale of securities by the central bank to restrict monetary expansion will be unsuccessful. But if there is a penal rate of discount, which is a rate higher than the market rate of interest, the banks will be reluctant to approach the central bank for additional financial help easily.

4. Banks Act Differently. Open market operations are successful only if the people also act the way the central bank expects them. When the central bank sells securities, it expects the business community and financial institutions to restrict the use of credit. If they simultaneously start dishoarding money, the act of selling securities by the central banks will not be a success in restricting credit. Similarly, the purchase of securities by the central bank will not be effective if people start hoarding money.

5. Pessimistic or Optimistic Attitude. Pessimistic or optimistic attitude of the business community also limits the operation of open market policy. When the central bank purchases securities and increases the supply of bank money, businessmen may be unwilling to take loans during a depression because of the prevailing pessimism among them. As aptly put by Crowther, "banks may place plenty of water before the public

horse, but the horse cannot be forced to drink, if it is afraid of loss through drinking water." On the other hand, if businessmen are optimistic during a boom, the sale of securities by the central bank to contract the supply of bank money and even the rise in market rates cannot discourage them from getting loans from the banks. On the whole, this policy is more successful in controlling booms than depressions.

6. Velocity of Credit Money not Constant. The success of open market operations depends upon a constant velocity of circulation of bank money. But the velocity of credit money is not constant. It increases during periods of brisk business activity and decreases in periods of falling prices. Thus a policy of contracting credit by the sale of securities by the central bank may not be successful by increased velocity of circulation of bank credit.

Conclusion. Despite these limitations, open market operations are more effective than the other instruments of credit control available with the central bank. This method is being successfully used for controlling credit in developed countries where the securities market is highly developed.

Open Market Operations vs Bank Rate Policy

The question arises whether the bank rate is more effective as an instrument of credit control or open market operations.

1. The bank rate policy influences the*cost and supply* of commercial bank credit, while open market operations affect the *cash reserves* of the commercial banks. Changes in the bank rate affect the credit creation power of the commercial banks only if they rediscount their bills with the central bank. If the banks do not feel the necessity of availing rediscounting facilities of the central bank, a rise in the bank rate will have no effect on the commercial banks. On the other hand, the lending power of the commercial banks is directly related to their cash reserves, and open market operations influence their cash reserves directly and immediately thereby affecting their credit creation power.
2. "From the standpoint of their strategic value to the central bank, open market operations possess a degree of superiority over

rediscount policy because of the fact that initiative is in the hands of the monetary authority in the case of the former, whereas bank rate policy is passive in the sense that its effectiveness depends on the responses of commercial banks and their customers to changes in bank rates."[2]

3. Open market operations are*flexible* with respect to timing and magnitude as compared with the bank rate policy. They are carried on continuously and do not have any destablilising effects on the economy that accompany changes in the bank rate.
4. It is further argued that since bank rate changes have destabilising effects on the economy, this policy should be used to correct permanent maladjustments in the money market rather than temporary maladjustments. On the other hand, open market policy can be used for correcting both temporary and permanent maladjustments in the money market.

Conclusion. But the experience of developed countries like the USA and the UK tells us that these two policies are not competitive but complementary to each other. If they are supplemented, they can control credit more effectively than individually. For instance, if the central bank raises the discount rate for the purpose of contracting credit, it will not be effective when the commercial banks have large excess reserves with them. They will continue to expand credit irrespective of the rise in the bank rate. But if the central bank first draws away to itself the excess reserves of the commercial banks by the sale of securities and then raises the bank rate, it will have the effect of contracting credit. Similarly, the sale of securities alone will not be so effective in contracting credit unless the bank rate is also raised. The sale of securities by the central bank will reduce the cash reserves of commercial banks but if the discount rate is low, the latter will get funds from the "discount window" of the central bank. So for an effective policy of credit control, bank rate policy and open market operations should be judiciously supplemented.

3. Variable Reserve Ratio

Variable reserve ratio (or required reserve ratio or legal minimum requirements), as a method of credit control was first suggested by Keynes

in his *Treatise on Money* (1930) and was adopted by the Federal Reserve System of the United States in 1935.

Every commercial bank is required by law to maintain a minimum percentage of its deposits with the central bank. The minimum amount of reserve with the central bank may be either a percentage of its time and demand deposits separately or of total deposits. Whatever the amount of money remains with the commercial bank over and above these minimum reserves is known as the excess reserves. It is on the basis of these excess reserves that the commercial bank is able to create credit. The larger the size of the excess reserves, the greater is the power of a bank to create credit, and vice versa. It can also be said that the larger the required reserve ratio, the lower the power of a bank to create credit, and vice versa.

When the central bank raises the reserve ratio of the commercial banks, it means that the latter are required to keep more money with the former. Consequently, the excess reserves with the commercial banks are reduced and they can lend less than before. On the contrary, if the central bank wants to expand credit, it lowers the reserve ratio so as to increase the credit creation power of the commercial banks. Thus by varying the reserve ratio of the commercial banks the central bank influences their power of credit creation and thereby controls credit in the economy.

2.G.N. Halm, *Monetary Theory*, p.59.

Limitations of Variable Reserve Ratio

The variable reserve ratio as a method of credit control has a number of limitations.

1. Excess Reserves. The commercial banks usually possess large excessive reserves which make the policy of variable reserve ratio ineffective. When the banks keep excessive reserves, an increase in the reserve ratio will not affect their lending operations. They will stick to the legal minimum requirements of cash to deposits and at the same time continue to create credit on the strength of the excessive reserves.

2. Clumsy Method. It is a clumsy method of credit control as compared with open market operations. This is because it lacks definiteness in the sense that it is inexact and uncertain as regards changes not only in the amounts of reserves but also the place where these changes can be made effective. It is not possible to tell "how much of active or potential reserve base" has been affected by changes in the reserve ratio. Moreover, the changes in reserves involve far larger sums than in the case of open market operations.

3. Discriminatory. It is discriminatory and affects different banks differently. A rise in the required reserve ratio will not affect those banks which have large excess reserves. On the other hand, it will hit hard the banks with little or no excess reserves. This policy is also discriminatory in the sense that non-banking financial intermediaries like co-operative societies, insurance companies, building societies, development banks, etc. are not affected by variations in reserve requirements, though they compete with the commercial banks for lending purposes.

4. Inflexible. This policy is inflexible because the minimum reserve ratio fixed by the central bank is applicable to banks located in all regions of the country. More credit may be needed in one region where there is monetary stringency, and it may be superfluous in the other region. Raising the reserve ratio for all banks is not justified in the former region though it is appropriate for the latter region.

5. Business Climate. The success of the method of credit control also depends on the business climate in the economy. If the businessmen are pessimistic about the future, as under a depression, even a sizable lowering of the reserve ratio will not encourage them to ask for loan. Similarly, if they are optimistic about profit expectations, a considerable rise in the variable ratio will not prevent them from asking for more loans from the banks.

6. Stability of Reserve Ratio. The effectiveness of this technique depends upon the degree of stability of the reserve ratio. If the commercial banks are authorised to keep widely fluctuating ratio, say between 10 per cent to 17 per cent, and a change in the upper or lower limit will have no effect on the credit creation power of the banks.

7. Other Factors. The reserve ratio held by the commercial banks is determined not only by legal requirements but also by how much they want to hold in relation to their deposits in addition to such requirements. This, in turn, will depend upon their expectations about future developments, their competition with other banks, and so on.

8. Depressive Effect. The variable reserve ratio has been criticised for exercising a depressive effect on the securities market. When the central bank suddenly directs the commercial banks to increase their reserve ratios, they may be forced to sell securities to maintain that ratio. This widespread selling of securities will bring down the prices of securities and may even lead to an utter collapse of the bond market.

9. Rigid. It is rigid in its operations because it does not distinguish between desired and undesired credit flows and can affect them equally.

10. Not for Small Changes. This method is more like an axe than a scalpel. It cannot be used for day-to-day and week-to-week adjustments but can be used to bring about large changes in the reserve positions of the commercial banks. Thus it cannot help in 'fine tuning' of the money and credit systems by making small changes.

Conclusion. The variable reserve ratios as a method of credit control is a very delicate and sensitive tool which not only produces a state of uncertainty among the banks but also adversely affects their liquidity and profitability. Therefore, according to De Kock, "it should be used with moderation and discretion and only under obvious abnormal conditions."

Variable Reserve Ratio vs Open Market Operations

There are divergent views about the superiority of variable reserve ratio over open market operations. To those who consider the former as a superior instrument of credit control, it is "a battery of the most improved type that a central bank can add to its armoury." They give the following arguments.

1. The variable reserve ratio affects the power of credit creation of the commercial banks more*directly*, *immediately*, and *simultaneously*

than open market operations. The central bank has simply to make a declaration for changing the reserve requirements of the banks and they have to implement it immediately. But open market operations require sale or purchase of securities which is a time consuming process.

2. When the central bank sells securities to the banks to control inflation, they are forced to buy them. They are, therefore, prevented from giving more loans to the private credit market. On the other hand, if the reserve ratio is raised, the banks will be required to keep larger balances with the central bank. They will also be faced with reduced earnings. They will, therefore, be induced to sell government securities and give more loans to the private credit market. Thus open market operations are more *effective* for controlling inflation than the change in reserve ratio.
3. Open market operations are superior to variations in the reserve ratio because they influence non-banking financial institutions. In every country there are non-banking financial intermediaries which deal in securities, bonds, etc. and also advance loans and accept deposits from the public. But they are outside the legal control of the central bank. Since they also deal in government securities, open market sale and purchase of such securities by the central bank also affect their liquidity position. But they are not required to keep any reserves with the central bank, unlike the commercial banks.
4. Variations in the reserve ratio are meant for making major and long-run adjustments in the*liquidity position* of the commercial banks. They are, therefore, not suited for making short-run adjustments in the volume of available bank reserves, as are done under open market operations.
5. The effectiveness of open market operations depends upon the existence of a *broad* and *well-organised market for securities*. Thus this instrument of credit control cannot be operative in countries which lack such a market. On the other hand, the method of variable reserve ratio does not require any such market for its operation and is applicable equally in developed and underdeveloped markets, and is thus superior to open market operations.
6. Since open market operations involve the sale and purchase of securities on a day-to-day and week-to-week basis, the commercial

banks and the central bank which deal in them are likely to incur *losses*. But variations in the reserve ratio do not involve any losses.

7. As a technique, reserve ratio can only influence the*volume of reserves* of the commercial banks. On the other hand, open market operations can influence not only the reserves of the commercial banks but also the pattern of the interest rate structure. Thus open market operations are more effective in influencing the reserves and the credit creation power of the banks than variations in reserve ratio.
8. The technique of variations in reserve ratio is*clumsy*, *inflexible*, and *discriminatory* whereas that of open market operations is *simple*, *flexible* and *non-discriminatory* in its effects.

Conclusion. It can be concluded from the above discussion of the relative merits and demerits of the two techniques that in order to have the best of the two, they should be used jointly rather than independently. If the central bank raises the reserve ratio, it should simultaneously start purchasing, and not selling, securities in those areas of the country where there is monetary stringency. On the contrary, when the central bank lowers the reserve requirments of the banks, it should also sell securities to those banks which already have excess reserves with them, and have been engaged in excessive lending. The joint application of the two policies will not be contradictory but complementary to each other.

Selective Credit Controls or Qualitative Methods

Selective or qualitative methods of credit control are meant to regulate and control the supply of credit among its possible users and uses. They are different from quantitative or general methods which aim at controlling the cost and quantity of credit. Unlike the general instruments, selective instruments do not affect the total amount of credit but the amount that is put to use in a particular sector of the economy. The aim of selective credit control is to channelise the flow of bank credit from speculative and other undesirable purposes to socially desirable and economically useful uses. They also restrict the demand for money by laying down certain conditions for borrowers. They therefore, embody the view that the monopoly of credit should in fact become a discriminating monopoly.

Prof. Chandler defines selective credit controls as those measures "that would influence the allocation of credit, at least to the point of decreasing the volume of credit used for selected purposes without the necessity of decreasing the supply and raising the cost of credit for all purposes." We discuss below the main types of selective credit controls generally used by the central banks in different countries.

(A) Regulation of Margin Requirements

This method is employed to prevent excessive use of credit to purchase or carry securities by speculators. The central bank fixes minimum margin requirements on loans for purchasing or carrying securities. They are, in fact, the percentage of the value of the security that cannot be borrowed or lent. In other words, it is the maximum value of loan which a borrower can have from the banks on the basis of the security (or collateral). For example, if the central bank fixes a 10 per cent margin on the value of a security worth Rs 1,000, then the commercial bank can lend only Rs 900 to the holder of the security and keep Rs 100 with it. If the central bank raises the margin to 25 per cent, the bank can lend only Rs 750 against a security of Rs 1,000. If the central bank wants to curb speculative activities, it will raise the margin requirements. On the other hand, if it wants to expand credit, it reduces the margin requirements.

Its Merits. This method of selective credit control has certain merits which make it unique.

1. It is non-discriminatory because it applies equally to borrowers and lenders. Thus it limits both the supply and demand for credit simultaneously.

2. It is equally applicable to commercial banks and non-banking financial intermediaries.

3. It increases the supply of credit for more productive uses.

4. It is a very effective anti-inflationary device because it controls the expansion of credit in those sectors of the economy which breed inflation.

5. It is simple and easy to administer since this device is meant to regulate the use of credit for specific purposes.

But the success of this technique requires that there are no leakages of bank credit for non-purpose loans to speculators.

Its Weaknesses. However, a number of weaknesses have appeared in this method over the years.

1. A borrower may not show any intention of purchasing stocks with his borrowed funds and pledge other assets as security for the loan. But it may purchase stocks through some other source.

2. The borrower may purchase stocks with cash which he would normally use to purchase materials and supplies and then borrow money to finance the materials and supplies already purchased, pledging the stocks he already has as security for the loan.

3. Lenders, other than commercial banks and brokers, who are not subject to margin requirements, may increase their security loans when commercial banks and brokers are being controlled by high margin requirements. Further, some of these nonregulated lenders may be getting the funds they lent to finance the purchase of securities from commercial banks themselves.

Despite these weaknesses in practice, margin requirements are a useful device of credit control.

(B) Regulation of Consumer Credit

This is another method of selective credit control which aims at the regulation of consumer instalment credit or hire-purchase finance. The main objective of this instrument is to regulate the demand for durable consumer goods in the interest of economic stability. The central bank regulates the use of bank credit by consumers in order to buy durable consumer goods on instalments and hire-purchase. For this purpose, it employs two devices: minimum down payments, and maximum periods of repayment. Suppose a bicycle costs Rs 500 and credit is available from the

commercial bank for its purchase. The central bank may fix the mimimum down payment to 50 per cent of the price, and the maximum period of repayment to 10 months. So Rs 250 will be the minimum which the consumer will have to pay to the bank at the time of purchase of the bicycle and the remaining amount in ten equal instalments of Rs 25 each. This facility will create demand for bicycles. The bicycle industry would expand along with the related industries such as tyres, tubes, spare parts, etc. and thus lead to inflationary situation in this and other sectors of the economy. To control it, the central bank raises the minimum down payment to 70 per cent and the maximum period of repayment to three instalements. So the buyer of a bicycle will have to pay Rs 350 in the beginning and remaining amount in three instalments of Rs 50 each. Thus if the central bank finds slump in particular industries of the economy, it reduces the amount of down payments and increases the maximum periods of repayment. Reducing the down payments tends to increase the demand for credit for particular durable consumer goods on which the central bank regulation is applied. Increasing the maximum period of repayment, which reduces monthly payments, tends to increase the demand for loans, thereby encouraging consumer credit. On the other hand, the central bank raises the amount of down payments and reduces the maximum periods of repayment in boom.

Its Merits. The regulation of consumer credit is more effective in controlling credit in the case of durable consumer goods during both booms and slumps, whereas general credit controls fail in this area. The reason is that the latter operate with a time lag. Moreover, the demand for consumer credit in the case of durable consumer goods is interest inelastic. Consumers are motivated to buy such goods under the influence of the demonstration effect and the rate of interest has little consideration for them.

Its Weaknesses. But this instrument has the following drawbacks:

1. It is cumbersome, technically defective and difficult to administer because it has a narrow base. In other words, it is applicable to a particular class of borrowers whose demand for credit forms an insignificant part of the total credit requirements. It, therefore,

discriminates between different types of borrowers.

2. This method affects only persons with limited incomes and leaves out higher income groups.
3. It tends to malallocate resources by shifting them away from industries which are covered by credit regulations and lead to the expansion of other industries which do not have any credit restrictions.

(C) Rationing of Credit

Rationing of credit is another selective method of controlling and regulating the purpose for which credit is granted by the commercial banks. It is generally of two types. The *first* is the variable portfolio ceiling. According to this method, the central bank fixes a ceiling on the aggregate portfolios of the commercial banks and they cannot advance loans beyond this ceiling. The *second* method is known as the variable capital assets ratio. This is the ratio which the central bank fixes in relation to the capital of a commercial bank to its total assets. In keeping with the economic exigencies, the central bank may raise or lower the portfolio ceiling, and also vary the capital assets ratio.

Rationing of credit has been used very effectively in Russia and Mexico. It is, therefore, 'a logical concomitant of the intensive and extensive planning adopted in regimented economies.' The technique also involves discrimination against larger banks because it restricts their lending power more than the smaller banks. Lastly, by rationing of credit for selective purposes, the central bank ceases to be the lender of the last resort. Therefore, central banks in mixed economies do not use this technique except under extreme inflationary situations and emergencies.

(D) Direct Action

Central banks in all countries frequently resort to direction action against commercial banks. Direction action is in the form of "directives" issued from time to time to the commercial banks to follow a particular policy which the central bank wants to enforce immediately. This policy may not be used against all banks but against erring banks. For example, the central bank refuses rediscounting facilities to certain banks which may be

granting too much credit for speculative purposes, or in excess of their capital and reserves, or restraints them from granting advances against the collateral of certain commodities, etc. It may also charge a penal rate of interest from those banks which want to borrow from it beyond the prescribed limit. The central bank may even threaten a commercial bank to be taken over by it in case it fails to follow its policies and instructions.

But this method of credit suffers from several limitations which have been enumerated by De Kock as "the difficulty for both central and commercial banks to make clear-cut distinctions at all times and in all cases between essential and non-essential industries, productive and unproductive activities, investment and speculation, or between legitimate and excessive speculation or consumption; the further difficulty of controlling the ultimate use of credit by second, third or fourth parties; the dangers involved in the division of responsibility between the central bank and commercial bank for the soundness of the lending operations of the latter; and the possibility of forfeiting the whole-hearted and active co-operations of the commercial banks as a result of undue control and intervention."[3]

(E) Moral Suasion

Moral suasion is the method of persuasion, of request, of informal suggestion, and of advice to the commercial banks which is usually adopted by the central bank. The executive head of the central bank calls a meeting of the heads of the commercial banks wherein he explains them the need for the adoption of a particular monetary policy in the context of the current economic situation, and then appeals to them to follow it. This "jawbone control" or "slaps on the wrist" method has been found to be highly effective as a selective method of credit control in India, New Zealand, Canada and Australia, though it failed in the USA.

Its Limitations. Moral suasion is a method "without any teeth" and hence its effectiveness is limited.

1. Its success depends upon the extent to which the commercial banks accept the central bank as their leader and need accommodation from it.
2. If the banks possess excessive reserves they may not follow the

adivce of the central bank, as is the case with the commercial banks in the USA.

3. Further, moral suasion may not be successful during booms and depressions when the economy is passing through waves of optimism and pessimism respectively. The bank may not pay heed to the advice of the central bank in such a situation.
4. In fact, moral suasion is not a control device at all, as it involves cooperation by the commercial banks rather than their coercion.

It may, however, be a success where the central bank commands prestige on the strength of the wide statutory powers vested in it by the government of the country.

(F) Publicity

The central bank also uses publicity as an instrument of credit control. It publishes weekly or monthly statements of the assets and liabilities of the commercial bank for the information of the public. It also publishes statistical data relating to money supply, prices, production and employment, and of capital and money market, etc. This is another way of exerting moral pressure on the commercial banks. The aim is to make the public aware of the policies being adopted by the commercial banks vis-a-vis the central bank in the light of the prevailing economic conditions in the country.

---

3. De Kock, *op. cit*., pp. 228-229.

It cannot be said with definiteness about the success of this method. It presupposes the existence of an educated and knowledgeable public about the monetary phenomena. But even in advanced countries, the percentage of such persons is negligible. It is, therefore, highly doubtful if they can exert any moral pressure on the banks to strictly follow the policies of the central bank. Hence, publicity as an instrument of selective credit control is only of academic interest.

Limitations of Selective Credit Controls

Though regarded superior to quantitative credit controls, yet selective credit controls are not free from certain limitations.

1. Limited Coverage. Like general credit controls, selective credit controls have a limited coverage. They are only applicable to the commercial banks but not to non-banking financial institutions. But in the case of the regulation of consumer credit which is applicable both to banking and non-banking institutions, it becomes cumbersome to administer this technique.

2. No Specificity. Selective credit controls fail to fulfil the specificity function. There is no guarantee that the bank loans would be used for the specific purpose for which they are sanctioned.

3. Difficult to distinguish between Essential and Non-essential Factors. It may be difficult for the central bank to distinguish precisely between essential and non-essential sectors and between speculative and productive investments for the purpose of enforcing selective credit controls. The same reasoning applies to the commercial banks for the purpose of advancing loans unless they are specifically laid down by the central bank.

4. Require Large Staff. The commercial banks, for the purpose of earning large profits, may advance loans for purposes other than laid down by the central bank. This is particularly so if the central bank does not have a large staff to check minutely the accounts of the commercial banks. As a matter of fact, no central bank can afford to check their accounts. Hence selective credit controls are liable to be ineffective in the case of unscrupulous banks.

5. Discriminatory. Selective controls unnecessarily restrict the freedom of borrowers and lenders. They also discriminate between different types of borrowers and banks. Often small borrowers and small banks are hit harder by selective controls than big borrowers and large banks.

6. Malallocation of Resources. Selective credit controls also lead to malallocation of resources when they are applied to selective sectors, areas and industries while leaving others to operate freely. They place

undue restrictions on the freedom of the former and affect their production.

7. Not Successful in Unit Banking. Unit banks being independent one-office banks in the USA operate on a small scale in small towns and meet the financial needs of the local people. Such banks are not affected by the selective credit controls of the Federal Reserve System because they are able to finance their activities by borrowing from big banks. So this policy is not effective in unit banking.

Conclusion. From the above discussion, it should not be concluded that selective credit controls are used to the total exclusion of general credit controls. Their demerits primarily arise from this thinking. In fact, they are an adjunct to general quantitative controls. They are meant to supplement the latter and are regarded only as a "second-line instrument".

EXERCISES

1. Discuss the essential functions of a central bank.
2. Distinguish between quantitative and qualitative methods of credit control.
3. Account for the recent shift from conventional quantitative control to qualitative control.
4. Between Bank Rate and Open Market Operations which is more effective as an instrument of credit control? Give reasons.
5. Explain the significance of selective credit controls. How do they operate and with what success ?

CHAPTER 31

THE MONETARIST REVOLUTION

MEANING

The "monetarist revolution" refers to the new and important contributions made to monetary theory and policy by Prof. Friedman and his colleagues at the University of Chicago. It was a sort of revolution against the views of Keynesians who held the view that "money does not matter." The Keynesians regarded the money supply as a passive factor in the economic system whose economic effects were highly unpredictable. On the other hand, in the monetarist revolution "only money matters" for three reasons: *one,* because the quantity of money is capable of being controlled fairly accurately by deliberate policy; *two*, because changes in the quantity of money can produce substantial changes in the flow of income, prices and other important variables; and *three,* because the relationships between stock of money and other assets are relatively stable and dependable."[1]

1. M. Friedman and D. Meiselman, in *Stabilisation Policies,* 1963.

MAIN FEATURES

The monetarist revolution possesses the following characteristics:

1. Money Supply Crucial Determinant

The money supply is the crucial determinant of economic activity in the short-run. It is the money supply that determines total spending, and therefore, output, employment and the price level. Thus there is a direct

link between the money supply and the national income. That link is the constant velocity of money. The constant velocity is expressed as *Y/M.* As a result of the stability of monetary velocity, a change in the money supply will change total spending and national income by a predictable amount. The demand for money is a stable function of income. The demand for money is the transactions demand for money which is determined by the level of income. If the central bank increases the money supply by purchasing securities, people who sell securities find that their holdings of money have increased. They will, therefore, spend their excess holdings of money partly on assets and partly on consumer goods and services. This spending will reduce their money balances and at the same time raise the national income. On the contrary, a reduction in the money supply by selling securities on the part of the central bank will reduce money holdings of the buyers of securities. They will, therefore, increase their money holdings partly by selling their assets and partly by reducing their consumption expenditure on goods and services. This will reduce the national income. Thus, on both counts, the demand for money remains stable.

According to Friedman, a change in the money supply causes a proportionate change in the price level or income or both. Given the demand for money, if the economy is operating at less than full employment level, an increase in the money supply will raise output and employment with a rise in total expenditure in the shortrun. An increase in the money supply causes national income to rise because with excess money supply, people start spending more until the demand and supply of money are equal. Thus an increase in the money supply in the short run raises output, employment and income. But a rise in the money supply in the longrun, with further increase in demand, prices and wages will increase. In the expectation of inflation, price-wage spiral will rise further. Thus there will be inflation due to inappropriate increase in money supply. That is why the monetarists regard inflation as a purely monetary phenomenon.

The monetarists regard their viewpoint as revolutionary as against the Keynesians because the rate of interest plays no part in influencing either the demand for money or the supply of money. Moreover, changes in the

money supply influence economic activity directly and not indirectly through changes in the interest rate like the Keynesians.

2. Transmission Mechanism of Monetary Influences

The transmission mechanism of monetary influences on economic activity involves reshuffling of both financial and real assets in the portfolios of economic units. Keynes considered only two types of assets, bonds and speculative cash balances, in his transmission mechanism. According to the monetarists, when the central bank increases the money supply by purchasing securities in the open market, their prices rise but yields fall due to a fall in the market rate of interest. People will, therefore, sell securities, and their holdings of money will increase. This raises the demand for financial and real assets. This will lead to the substitution of excess money balances for financial assets and durable consumer goods. The increase in aggregate expenditure on assets and goods will tend to raise the national income. This is the *substitution effect* of the portfolio adjustment process.

Further, when money is converted into securities with their purchase in the open market, the nominal wealth of the community increases. This is the *direct wealth effect.* Again, with the fall in the market rate of interest, the market value of current capital stock increases which also raises the nominal wealth of the society. This is the *interest induced effect* of wealth effect. But these effects increase the net wealth. Consequently, people buy financial and real assets which lead to the production of new producer and consumer durables goods and encourage the purchase of consumer non-durable products.

In the monetarist system, a central bank cannot influence interest rates through changes in the supply of money. If it tries to reduce the interest rate by releasing large quantity of money, this will only cause inflation since the economy is already at near-full employment level. In fact, this will reduce the real money supply and increase interest rates. The monetarists consider this as revolutionary.

3. Deteminant of Real National Income

In the long-run the level of real national income is determined by the forces of demand and supply. This is based on the assumptions that prices and wages in all markets are inherently flexible. They rise in response to excess demand and fall in response to excess supply. If some prices are inflexible, the burden of their adjustments would fall on other products. Thus the economy is usually at or near the full employment level where there is no involuntary unemployment. Friedman refers to this as the *natural rate of unemployment.* This view is in marked contrast to the Keynesian view that there is always underemployment equilibrium in the economy and unemployment is involuntary.

4. Stable Economy

The monetarists hold that the economy is stable. They do not believe like the Keynesians that it is subject to wide or sudden fluctuations due to changes in the propensities to consume and invest. According to them, instability exists in the economy on account of monetary and fiscal policies adopted by the government. These policies destabilise rather than stabilise the economy. Friedman does not favour even countercyclical monetary policy. According to him, monetary policy might do more harm than good because of the operation lag. On the average, it takes a long time for a change in the money supply to affect national income. The time lag involved is so large that countercyclical monetary policy might actually have a destabilising effect on the economy. Fiscal policy has no place in the monetarist system. It does not affect the economy unless it is accompanied by changes in the money supply. So there is no need for fiscal policy as the same results can be achieved by monetary policy. However, to stabilise the economy and avoid inflation, Friedman advocates a steady and inflexible growth in the rate of money supply. When the money supply increases at the same rate as output, the national income grows without inflation.

---

1. For a detailed analysis of the monetarist policy measures, refer to the last section of the chapter *Monetarism vs Keynesianism.*

2. Nicholas Kaldor, *Failure of Monetarism,* C.D. Deshmukh Memorial Lecture, RBI, 18 January 1984.

5. Important Role of Expectations

Expectations play an important role in the monetarist's view. Every person whether he is a businessman, consumer or worker is capable of correctly anticipating the effects of his own and other persons' actions. The monetarists hold that expectations are rational. "Decisions taken on the basis of such expectations will cause the anticipated future results to occur even more quickly, if not at once. Thus intelligent expectations are self-reinforcing and stabilising so long as the government does not create false signals by erratic and irrational intervention." They have revolutionised economic thinking through the rational expectations hypothesis. For example, the rational expectationists deny the possibility of any inflation-unemployment trade-off even in the short run. Economists regard the above views of the monetarists as revolutionary.[1]

CRITICISMS

Prof. Kaldor does not regard the above views of the monetarists as revolutionary. He characterises them as a "counter-reformation—the reaction against the new economics of the 1930s and return to 19th century orthodoxies."[2] He and many Keynesians have criticised the monetarist tenets on the following grounds:

1. Money Supply Endogenous

The supply of money is varied by the monetary authorities in an exogenous manner in Friedman's system. But the fact is that in the United States the money supply consists of bank deposits created by changes in bank lending. Bank lending, in turn, is based upon bank reserves which expand and contract with (a) deposits and withdrawals of currency by non-bank financial intermediaries; *(b)* borrowings by commercial banks from the Federal Reserve System; (c) inflows and outflows of money from and to abroad; and *(d)* purchase and sale of securities by the Federal Reserve System. The first three items definitely impart an endogenous element to the money supply. Thus the money supply is not exclusively exogenous, as assumed by Friedman. It is mostly endogenous.

3. Nicholos Kaldor, *The Scourge of Monetarism*, 1982.

2. Demand for Money not Stable

Regarding the stability of the demand for money, Prof. Kaldor found that the demand for money as a proportion of income is neither stable between countries nor stable over time except in some countries.

3. Money Supply and GNP not Positively Correlated

Money supply and money GNP have been found to be positively correlated in Friedman's findings. But Kaldor found his evidence to be largely irrelevant. For example, he found that in Switzerland, Italy and Japan, the money supply on the broad definition, M_3had been rising for over twenty years in relation to incomes, while it had been falling in the US and the UK. Even on the narrow definition, M_1 the money supply in Switzerland was nearly three times as great as in the UK or the US as a proportion of the GNP. From this, he concludes that "yet no one would regard Switzerland as an 'inflation prone' country (let alone more inflation prone) than the US or the UK."[3]

4. Neglects the Role of NBFIs

The transmission mechanism explained by the monetarists has also been questioned. The Radcliffe Committee and Gurley and Shaw criticise the monetarist transmission mechanism for neglecting the role played by the non-bank financial intermediaries and their effects on real and financial assets.

5. Real World does not Approximate to a General Equilibrium System

The monetarist view that prices in all markets are completely flexible, is based on the Walrasian general equilibrium model of the economy. This implies that the economy is at the full employment level. Critics point out that wages and prices do not adjust themselves simultaneously in the Walrasian sense. In fact, trade unions are engaged in wage bargains with the rise in prices in the past. Similarly, entrepreneurs try to adjust their reduced profit margins which have been eroded by inflation. This further

increases inflation. As pointed out by Kaldor, the monetarists failed to recognise the all-important difference between a demand-inflation and cost-inflation. Thus the real world does not approximate to a general equilibrium system.

6. Economy not Inherently Stable

The monetarists contend that the economy is inherently stable and it is interference in the form of monetary policy that brings instability. This view has not been accepted by the Keynesians who argue that there are frequent wild and erratic shocks in the economy due to variations in investment and consumption spendings that produce business cycles. This necessitates appropriate countercyclical monetary and fiscal policies.

7. Money Supply fails to grow at a Smooth and Steady Rate

Further, to stabilise the economy and avoid inflation, the monetarists favour a steady growth in the rate of money supply, and the rate of interest finds no place in their policy frame. But the experience of both the US and the UK tells a different story where the monetarist monetary policy was put into operation in 1979-80. “The money supply failed to grow at a smooth and steady rate; its behaviour exhibited a series of wriggles. The rate of interest and the rate of inflation, though both were very high at the start, soared to unprecedented heights in a very short time.”

Thus on the basis of above criticisms, it can be concluded that the monetarist viewpoint was not a revolution but a sort of reformation of the Keynesian economics and the revival of the orthodox monetarism.

EXERCISES

1. What is monetarist revolution? How has it affected our views about inflation and monetary policy ?

2. What is meant by monetarist revolution? Explain its main features.

3. Critically examine the main features of the monetarist revolution.

4. “The views of monetarists are a counter-reformation rather than revolutionary.” Examine this statement.

CHAPTER 32

THE DEMAND FOR MONEY

INTRODUCTION

The demand for money arises from two important functions of money. The first is that money acts as a medium of exchange and the second is that it is a store of value. Thus individuals and businesses wish to hold money partly in cash and partly in the form of assets.

What explains changes in the demand for money ? There are two views on this issue. The first is the "scale" view which is related to the impact of the income or wealth level upon the demand for money. The demand for money is directly related to the income level. The higher the income level, the greater will be the demand for money. The second is the "substitution" view which is related to relative attractiveness of assets that can be substituted for money. According to this view, when alternative assets like bonds become unattractive due to fall in interest rates, people prefer to keep their assets in cash, and the demand for money increases, and vice versa. The scale and substitution view combined together have been used to explain the nature of the demand for money which has been split into the transactions demand, the precautionary demand and the speculative demand. There are three approaches to the demand for money: the classical, the Keynesian, and the post-Keynesian.[1] We discuss these approaches below.

THE CLASSICAL APPROACH

The classical economists did not explicitly formulate demand for money

theory but their views are inherent in the quantity theory of money. They emphasized the transactions demand for money in terms of the velocity of circulation of money. This is because money acts as a medium of exchange and facilitates the exchange of goods and services. In Fisher's "Equation of Exchnage",

$$MV=PT$$

where M is the total quantity of money, V is its velocity of circulation, P is the price level, and T is the total amount of goods and services exchanged for money.

The right hand side of this equation PT represents the demand for money which, in fact, "depends upon the value of the transactions to be undertaken in the economy, and is equal to a constant fraction of those transactions." MV represents the supply of money which is given and in equilibrium equals the demand for money. Thus the equation becomes

$$MV= PT$$

This transactions demand for money, in turn, is determined by the level of full employment income. This is because the classicists believed in Say's Law whereby supply created its own demand, assuming the full employment level of income. Thus the demand for money in Fisher's approach is a constant proportion of the level of transactions, which in turn, bears a constant relationship to the level of national income. Further, the demand for money is linked to the volume of trade going on in an economy at any time. Thus its underlying assumption is that people hold money to buy goods.

But people also hold money for other reasons, such as to earn interest and to provide against unforeseen events. It is, therefore, not possible to say that V will remain constant when M is changed. The most important thing about money in Fisher's theory is that it is transferable. But it does not explain fully why people hold money. It does not clarify whether to include as money such items as time deposits or savings deposits that are not immediately available to pay debts without first being converted into currency.

It was the Cambridge cash balances approach which raised a further question: Why do people actually want to hold their assets in the form of money ? With larger incomes, people want to make larger volumes of transactions and that larger cash balances will, therefore, be demanded.

The Cambridge demand equation for money is

$$Md = kPY$$

where *Md* is the demand for money which must equal the supply of money (*Md=Ms*) in equilibrium in the economy. *k* is the fraction of the real money income (*PY*) which people wish to hold in cash and demand deposits or the ratio of money stock to income, *P* is the price level, and *Y* is the aggregate real income. This equation tells us that "other things being equal, the demand for money in normal terms would be proportional to the nominal level of income for each individual, and hence for the aggregate economy as well."

1.David E.W. Laidler, *The Demand for Money: Theories and Evidence*, 1972, Part Two.

Its Critical Evaluation

This approach includes time and saving deposits and other convertible funds in the demand for money. It also stresses the importance of factors that make money more or less useful, such as the costs of holding it, uncertainty about the future and so on. But it says little about the nature of the relationship that one expects to prevail between its variables, and it does not say too much about which ones might be important.

One of its major criticisms arises from the neglect of store of value function of money. The classicists emphasized only the medium of exchange function of money which simply acted as a go-between to facilitate buying and selling. For them, money performed a *neutral* role in the economy. It was barren and would not multiply, if stored in the form of wealth. This was an erroneous view because money performed the "asset" function when it is transformed into other forms of assets like bills, equities, debentures, real assets (houses, cars, TVs, and so on), etc. Thus

the neglect of the asset function of money was the major weakness of the classical approach to the demand for money which Keynes remedied.

THE KEYNESIAN APPROACH : LIQUIDITY PREFERENCE

Keynes in his *General Theory* used a new term "liquidity preference" for the demand for money. Keynes suggested three motives which led to the demand for money in an economy:(1) the transactions demand, (2) the precautionary demand, and (3) the speculative demand.

The Transactions Demand for Money

The transactions demand for money arises from the medium of exchange function of money in making regular payments for goods and services. According to Keynes, it relates to "the need of cash for the current transactions of personal and business exchange." It is further divided into income and business motives. The *income motive* is meant "to bridge the interval between the receipt of income and its disbursement." Similarly, the *business motive* is meant "to bridge the interval between the time of incurring business costs and that of the receipt of the sale proceeds." If the time between the incurring of expenditure and receipt of income is small, less cash will be held by the people for current transactions, and vice versa. There will, however, be changes in the transactions demand for money depending upon the expectations of income recipients and businessmen. They depend upon the level of income, the interest rate, the business turnover, the normal period between the receipt and disbursement of income, etc.

Given these factors, the transactions demand for money is a *direct proportional* and *positive function* of the level of income, and is expressed as

$$L_T=kY$$

where L_T is the transactions demand for money, k is the proportion of income which is kept for transactions purposes, and Y is the income.

This equation is illustrated in Figure 1 where the line *kY* represents a linear and proportional relation between transactions demand and the level of income. Assuming *k*=1/4 and income Rs 1000 crores, the demand for transactions balances would be Rs 250 crores, at point *A*. With the increase in income to Rs 1200 crores, the transactions demand would be Rs 300 crores at point *B* on the curve *kY*. If the transactions demand falls due to a change in the institutional and structural conditions of the economy, the value of *k* is reduced to say, 1/5, and the new transactions demand curve is *k'Y*. It shows that for income of Rs 1000 and 1200 crores, transactions balances would be Rs 200 and 240 crores at points *C* and *D* respectively in the figure. "Thus we conclude that the chief determinant of changes in the actual amount of the transactions balances held is changes in income. Changes in the transactions balances are the result of movements along a line like *kY* rather than changes in the slope of the line. In the equation, changes in transactions balances are the result of changes in *Y* rather than changes in *k*."[3]

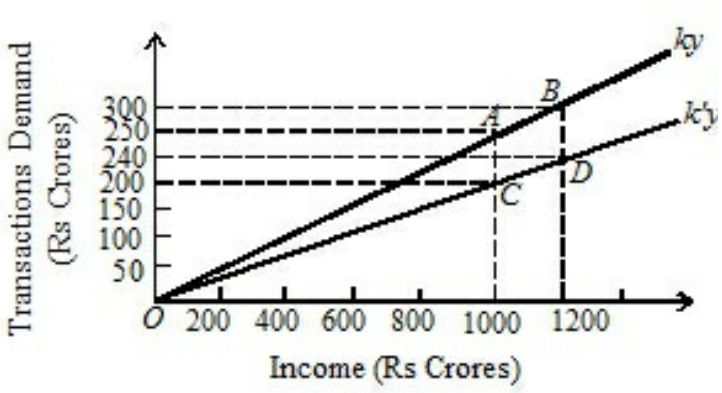


Fig. 1

Interest Rate and Transactions Demand. Regarding the rate of interest as the determinant of the transactions demand for money, Keynes made the L_T function *interest inelastic*. But he pointed out that the "demand for money in the active circulation is also to some extent a function of the rate of interest, since a higher rate of interest may lead to a more economical use of active balances."[4] "However, he did not stress the role of the rate of interest in this part of his analysis, and many of his popularizers ignored it altogether."[5] Two post-Keynesian economists William J. Baumol[6] and James Tobin[7] have shown that the rate of interst is an important determinant of transactions demand for money. The have also pointed out that the relationship between transactions demand for money and income is not linear and proportional. Rather, changes in income lead to proportionately smaller changes in transactions demand.

Transactions balances are held because income received once a month is not spent on the same day. In fact, an individual spreads his expenditure evenly over the month. Thus a portion of money meant for transactions purposes can be spent on short-term interest-yielding securities. It is possible to "put funds to work for a matter of days, weeks, or months in interest-bearing securities such as U.S. Treasury bills or commercial paper and other short-term money market instruments. The problem here is that there is a cost involved in buying and selling. One must weigh the financial cost and inconvenience of frequent entry to and exit from the market for securities against the apparent advantage of holding interest-bearing securities in place of idle transactions balances. Among other things, the cost per purchase and sale, the rate of interest, and the frequency of purchases and sales determine the profitability of switching from ideal transactions balances to earning assets. Nonetheless, with the cost per purchase and sale given, there is clearly some rate of interest at which it becomes profitable to switch what otherwise would be transactions balances into interest-bearing securities, even if the period for which these funds may be spared from transactions needs is measured only in weeks. The higher the interest rate, the larger will be the fraction of any given amount of transactions balances that can be profitably diverted into securities."[8]

---

3. Edward Shapiro, *op. cit*., pp. 367-68.

4. J.M.Keynes in A.D. Gayer (ed.), *The Lessons of Monetary Experience*, 1937, p. 149.

5. David E.W.Laidler, *op. cit*., p. 52.

6. William J.Baumol, "Transactions Demand for Cash--An Inventory Theoretic Approach", *Quarterly Journal of Economics*, November 1952.

7. James Tobin, "The Interest Elasticity of the Transactions Demand for Cash", *Review of Economics and Statistics*, August 1956.

The structure of cash and short-term bond holdings is shown in Figure 2 (A), (B) and (C). Suppose an individual receives Rs 1200 as income on the first of every month and spends it evenly over the month. The month has four weeks. His saving is zero. Accordingly, his transactions demand for money in each week is Rs 300. So he has Rs 900 idle money in the first week, Rs 600 in the second week, and Rs 300 in the third week. He

will, therefore, convert this idle money into interest-bearing bonds, as illustrated in Panel (B) and (C) of Figure 2. He keeps and spends Rs 300 during the first week (shown in Panel B), and invests Rs 900 in interest-bearing bonds (shown in Panel C). On the first day of the second week, he sells bonds worth Rs. 300 to cover cash transactions of the second week and his bond holdings are reduced to Rs 600. Similarly, he will sell bonds worth Rs 300 in the beginning of the third week and keep the remaining bonds amounting to Rs 300 which he will sell on the first day of the fourth week to meet his expenses for the last week of the month. The amount of cash held for transactions purposes by the individual during each week is shown in saw-tooth pattern in Panel (B), and the bond holdings in each week are shown in blocks in Panel (C) of Figure 2.

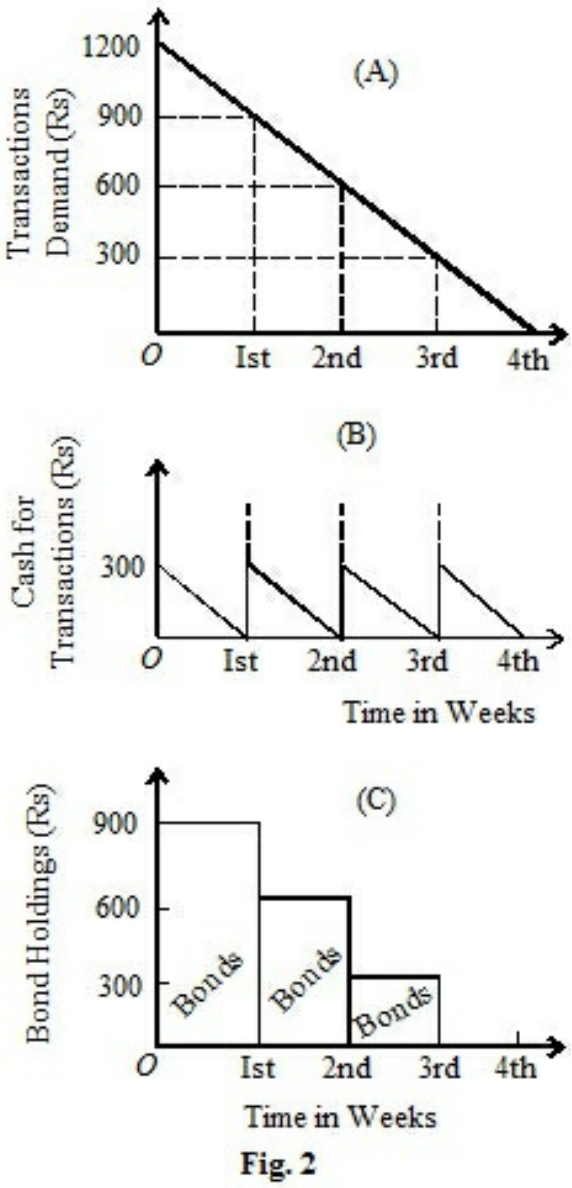


Fig. 2

8. Edward Shapiro, *op, cit*., pp. 368-69.

The modern view is that the transactions demand for money is a function of both income and interest rates which can be expressed as

$L_T = f(Y,r)$.

This relationship between income and interest rate and the transactions demand for money for the economy as a whole is illustrated in Figure 3. We saw above that $L_T = kY$. If Y=Rs 1200 crores and k=1/4, then L_T=Rs 300 crores.

This is shown as Y_1 curve in Figure 3. If the income level rises to Rs 1600 crores, the transactions demand also increases to Rs 400 crores, given k=1/4. Consequently, the transactions demand curve shifts to Y_2. The transactions demand curves Y_1 and Y_2 are interest-inelastic so long as the rate of interest does not rise above r_8 per cent. As the rate of interest starts rising above r_8, the transactions demand for money becomes interest elastic. It indicates that "given the cost of switching into and out of securities, an interest rate above 8 per cent is sufficiently high to attract some amount of transactions balances into securities." The backward slope of the Y_1 curve shows that at still higher rates, the transaction demand for money declines. Thus when the rate of interest rises to r_{12}, the transactions demand declines to Rs 250 crores with an income level of Rs 1200 crores. Similarly, when the national income is Rs 1600 crores, the transactions demand would decline to Rs 350 crores at r_{12} interest rate. Thus the transactions demand for money varies *directly* with the level of income and *inversely* with the rate of interest.

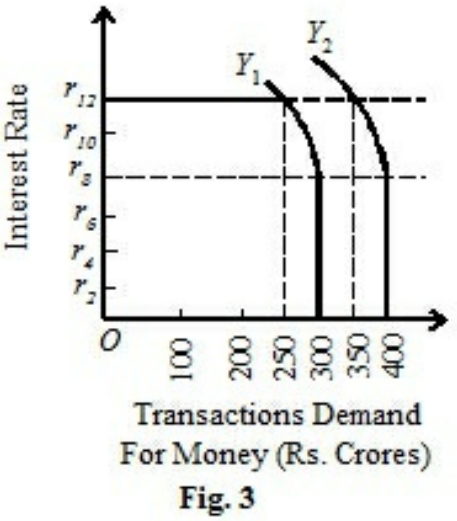


Fig. 3

The Precautionary Demand for Money

The precautionary motive relates to "the desire to provide for contingencies requiring sudden expenditures and for unforeseen opportunities of advantageous purchases." Both individuals and businessmen keep cash in reserve to meet unexpected needs. Individuals hold some cash to provide for illness, accidents, unemployment and other unforeseen contingencies. Similarly, businessmen keep cash in reserve to tide over unfavourable conditions or to gain from unexpected deals. Therefore, "money held under the precautionary motive is rather like water kept in reserve in a water tank." The precautionary demand for money depends upon the level of income, business activities, opportunities for unexpected profitable deals, availability of cash, the cost of holding liquid assets in bank reserves, etc.

Keynes held that the precautionary demand for money, like transactions demand, was a function of the level of income. But the post-Keynesian economists believe that like transactions demand, it is inversely related to high interest rates. The transactions and precautionary demand for money will be unstable, particularly if the economy is not at full employment level and transactions are, therefore, less than the maximum, and are liable to fluctuate up or down. Since precuationary demand, like transactions demand is a function of income and interest rates, the demand for money for these two purposes is expressed in the single equation $LT = f(Y,r)$.[9] Thus the precautionary demand for money can also be explained diagrammatically in terms of Figures 2 and 3.

The Speculative Demand for Money

The speculative (or asset or liquidity preference) demand for money is "for securing profit from knowing better than the market what the future will bring forth".Individuals and businessmen having funds, after keeping enough for transactions and precautionary purposes, like to make a speculative gain by investing in bonds. Money held for speculative purposes is a liquid store of value which can be invested at an opportune moment in interest-bearing bonds or securities.

Bond prices and the rate of interest are *inversely* related to each other. Low bond prices are indicative of high interest rates, and high bond prices reflect low interest rates. A bond carries a fixed rate of interest. For instance, if a bond of the value of Rs 100 carries 4 per cent interest and the market rate of interest rises to 8 per cent, the value of this bond falls to Rs 50 in the market. If the market rate of interest falls to 2 per cent, the value of the bond will rise to Rs 200 in the market.

This can be worked out with the help of the equation, $V = \frac{R}{r}$

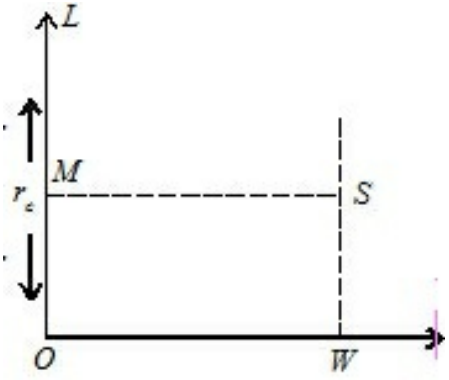
where V is the current market value of a bond, R is the annual return on the bond, and r is the rate of return currently earned or the market rate of interest. So a bond worth Rs 100 (V) and carrying a 4 per cent rate of interest (r), gets an annual return (R) of Rs 4, that is, V=Rs 4/0.04=Rs100. When the market rate of interest rises to 8 per cent, then V=Rs

4/0.08=Rs50; when it falls to 2 per cent, then *V*=Rs 4/0.02=Rs 200.

Thus individuals and businessmen can gain by buying bonds worth Rs 100 each at the market price of Rs 50 each when the rate of interest is high (8 per cent), and sell them again when they are dearer (Rs 200 each when the rate of interest falls (to 2 per cent).

According to Keynes, it is expectations about changes in bond prices or in the current market rate of interest that determine the speculative demand for money. In explaining the speculative demand for money, Keynes had a normal or critical rate of interest (r_c) in mind. If the current rate of interest (r) is above the "critical" rate of interest, businessmen expect it to fall and bond prices to rise. They will, therefore, buy bonds to sell them in future when their prices rise in order to gain thereby. At such times, the speculative demand for money would fall. Conversely, if the current rate of interest happens to be below the critical rate, businessmen expect it to rise and bond prices to fall. They will, therefore, sell bonds in the present if they have any, and the speculative demand for money would increase. Thus when $r>r_c$, an investor holds all his liquid assets in bonds, and when $r<r_c$ his entire holdings go into money. But when $r = r_c$, he becomes indifferent to hold bonds or money.

This relationship between an individual's demand for money and the rate of interest is shown in Figure 4 where the horizontal axis shows the individual's demand for money for speculative purposes and the current and critical interest rates on the vertical axis. The figure shows that when r is greater than r_c , the asset holder puts all his cash balances in bonds and his demand for money is zero. This is illustrated by the *LM* portion of the vertical axis. When r falls below r_c , the individual expects more capital losses on bonds as against the interest yield. He, therefore, converts his entire holdings into money, as shown by *OW* in the figure. This relationship between an individual asset holder's demand for money and the current rate of interest



Speculative Demand for Money

Fig. 4

gives the discontinuous step demand for money curve *LMSW*.

---

9. Keynes expressed the amount held under these two motives (M_1) as a function (L_1) of the level of income (Y), $M_1=L_1(Y)$. By making it a function of interest rate (r), it can be written as $M_1=L_1(Y,r)$.

For the economy as a whole the individual demand curve can be aggregated on this presumption that individual asset-holders differ in their critical rates r_c . It is a smooth curve which slopes downward from left to right, as shown in Figure 5.

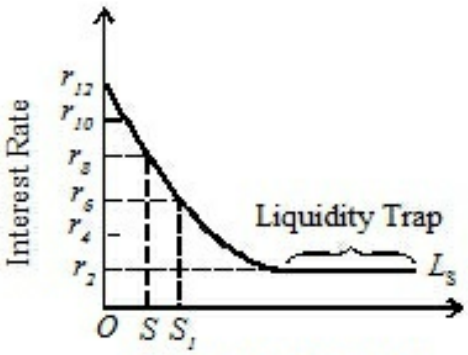


Fig. 5

Thus the speculative demand for money is a decreasing function of the rate of interest. The higher the rate of interest, the lower the speculative demand for moeny, and the lower the rate of interest, the higher the speculative demand for money. It can be expressed algebraically as $Ls = f(r)$, where Ls is the speculative demand for money and r is the rate of interest.[10] Geometrically, it is shows in Figure 5. The figure shows that at a very high rate of interest r_{12}, the speculative demand for money is zero and businessmen invest their cash holdings in bonds because they believe that the interest rate cannot rise further. As the rate of interest falls to say, r_8 the speculative demand for money is OS. With a further fall in the interest rate to r_6, it rises to OS_1. But at a very low rate of interest r_2, the L_S curve becomes perfectly elastic. This is know as the *liquidity trap* when people prefer to keep money in cash rather than invest in bonds and the speculative demand for money is infinitely elastic. Thus the shape of the Ls curve[11] shows that as the interest rate rises, the speculative demand for money declines, and with the fall in the interest rate, it increases. Thus the Keynesian speculative demand for money function is highly volatile, depending upon the behaviour of interest rates.

Note on Liquidity Trap

Keynes visualised conditions in which the speculative demand for money would be highly or even totally elastic so that changes in the quantity of money would be fully absorbed into speculative balances. This is the famous Keynesian *liquidity trap*. In this case, changes in the quantity of money have no effects at all on prices or income. According to Keynes, this is likely to happen when the market interest rate is very low so that yields on bonds, equities and other securities will also be low.

At a very low rate of interest, such as r_2, in Figure 5, the *Ls* curve becomes *perfectly elastic* and the speculative demand for money is infinitely elastic. This portion of the *Ls* curve is known as the liquidity trap. At such a low rate, people prefer to keep money in cash rather than invest in bonds because purchasing bonds will mean a definite loss. People will not buy bonds so long as the interest rate remains at the low level and they will be waiting for the rate of interest to return to the "normal" level and bond prices to fall.

---

10. Keynes expressed it as $M_2=L_2(r)$

11. This is also known as the liquidity preference curve, *LP*.

According to Keynes, as the rate of interest approaches zero, the risk of loss in holding bonds becomes greater. "When the price of bonds has been bid up so high that the rate of interest is, say, only 2 per cent or less, a very small decline in the price of bonds will wipe out the yield entirely and a slightly further decline would result in loss of the part of the principal." Thus the lower the interest rate, the smaller the earnings from bonds. Therefore, the greater the demand for cash holdings. Consequently, the *Ls* curve will become perfectly elastic.

Further, according to Keynes, "a long-term rate of interest of 2 per cent leaves more to fear than to hope, and offers, at the same time, a running yield which is only sufficient to offset a very small measure of fear." This makes the *Ls* curve "virtually absolute in the sense that almost everybody prefers cash to holding a debt which yields so low a rate of interest."

Prof. Modigliani believes that an infinitely elastic *Ls* curve is possible in a period of great uncertainty when price reductions are anticipated and the

tendency to invest in bonds decreases, or if there prevails "a real scarcity of investment outlets that are profitable at rates of interest higher than the institutional minimum."[12]

Its Implications. The phenomenon of liquidity trap possesses certain important implications:

First, the monetary authority cannot influence the rate of interest even by following a cheap money policy. An increase in the quantity of money cannot lead to a further decline in the rate of interest in a liquidity trap situation.

Second, the rate of interest cannot fall to zero.

Third, the policy of a general wage cut cannot be efficacious in the face of a perfectly elastic liquidity preference curve, such as *Ls* in Figure 5. No doubt, a policy of general wage cut would lower wages and prices, and thus release money from transactions to speculative purpose, the rate of interest would remain unaffected because people would hold money due to the prevalent uncertainty in the money market.

Last, if new money is created, it instantly goes into speculative balances and is put into bank vaults or cash boxes instead of being invested. Thus there is no effect on income. Income can change without any change in the quantity of money. Thus monetary changes have a weak effect on economic activity under conditions of absolute liquidity preference.

The Total Demand for Money

According to Keynes, money held for transactions and precautionary purposes is primarily a function of the level of income, $L_T = f(Y)$, and the speculative demand for money is a function of the rate of interest, $Ls = f(r)$. Thus the total demand for money is a function of both income and the interest rate :

$$L_T + L_s = f(Y) + f(r)$$

$$L_T + L_S = f(Y) + f(r)$$

or $$L = f(Y) + f(r)$$

or $$L = f(Y, r)$$

where L represents the total demand for money.

---

12. F. Modigliani, "Liquidity Preference and the Theory of Interest and Money," *Economica* January 1944.

Thus the total demand for money can be derived by the lateral summation of the demand function for transactions and precautionary purposes and the demand function for speculative purposes, as illustrated in Figure 6 (A), (B) and (C)*. Panel (A) of the Figure shows *OT*, the transactions and precautionary demand for money at *Y* level of income and different rates of interest. Panel (B) shows the speculative demand for money at various rates of interest. It is

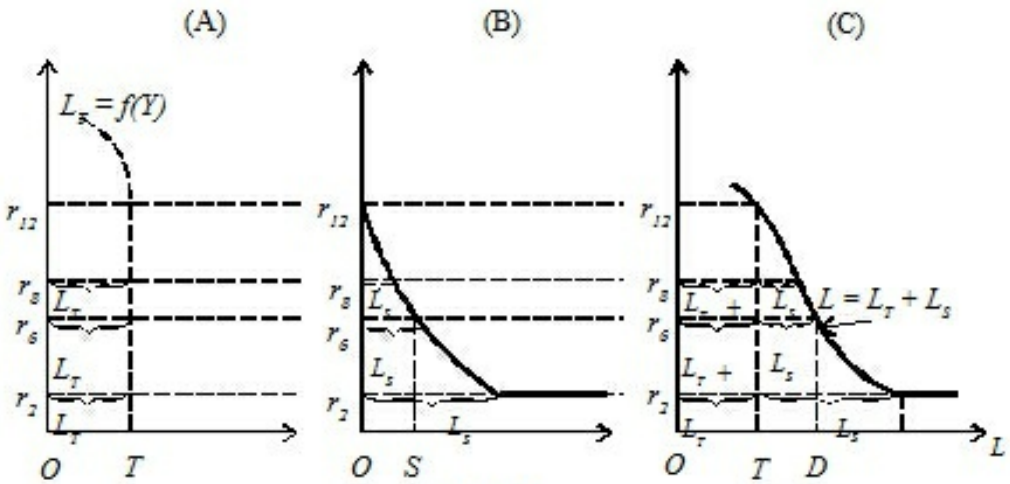


Fig. 6

an inverse function of the rate of interest. For instance, at r_6 rate of interest it is *OS* and as the rate of interest falls to r_2, the *Ls* curve becomes perfectly elastic. Panel (C) shows the total demand curve for money *L* which is a lateral summation of L_T and L_S curves : $L=L_T+L_S$. For example, at r_6 rate of interest, the total demand for money is *OD* which is the sum of transactions and precautionary demand *OT plus* the speculative demand *TD*, *OD=OT+TD*, where *TD = OS*. At r_2 interest rate, the total demand for

money curve also becomes perfectly elastic, showing the position of liquidity trap.

THE POST-KEYNESIAN APPROACHES

Keynes believed that the transactions demand for money was primarily interest inelastic. Prof. Baumol has analysed the interest elasticity of the transactions demand for money on the basis of his inventory theroretical approach. Further, in the Keynesian analysis the speculative demand for money is analysed in relation to uncertainty in the market. Prof. Tobin has given an alternative theory which explains liquidity preference as behaviour towards risk. The two approaches to the liquidity preference theory are discussed below.

---

*Adapted from Edward shapiro,*op,cit.*,

1. Baumol's Inventory Theoretic Approach[13]

William Baumol has made an important addition to the Keynesian transactions demand for money. Keynes regarded transactions demand for money as a function of the level of income, and the relationship between transactions demand and income as linear and proportional. Baumol shows that the relation between transactions demand and income is neither linear nor proportional. Rather, changes in income lead to less than proportionate changes in the transactions demand for money. Further, Keynes considered transactions demand as primarily interest inelastic. But Baumol analyses the interest elasticity of the transactions demand for money.

Its Assumptions

Baumol's theory is based on the following assumptions:

1. The transactions between money and bonds are transparent and occur in a steady stream.
2. The bond market is perfect where there is easy conversion of bonds into cash and vice versa.

3. There is a fixed cost in exchanging bonds for cash and vice versa.
4. The holding of cash involves interest cost and non-interest costs.
5. The interest cost (or rate of interest) is constant over the year.
6. The non-interest costs such as brokerage fee, mailing expenses, etc. are also fixed over the year.

The Theory

Given these assumptions, Baumol's analysis is based on the holding of an optimum inventory of money for transactions purposes by a firm or an individual. He writes: "A firm's cash balance can usually be interpreted as an inventory of money which its holder stands ready to exchange against purchase of labour, raw materials, etc." Cash balances are held because income and expenditure do not take place simultaneously. "But it is expensive to tie up large amounts of capital in the form of cash balances. For that money could otherwise be used profitably elsewhere in the firm or it could be invested profitably in securities." Thus the alternative to holding cash balances is bonds which earn interest. A firm would always try to keep minimum transactions balances in order to earn maximum interest from its assets. The higher the interest rate on bonds, the lesser the transactions balances which a firm holds.

Baumol assumes that a firm receives Y dollars once per time period, say a year, which are spent at a constant rate over the period. It is, therefore, always profitable for the firm to spend idle funds on buying bonds which can be sold when it needs cash for transactions purposes.

The structure of cash for holdings and bond holdings by a firm is shown in Figure 7. Suppose the firm has $ 1,200 which it has to spend every quarter at a constant rate over the year. Out of this, it keeps $ 400 in cash for transactions purposes and buys bonds with the remaining amount of $800. Half the bonds purchased carry maturity of $1/3t$ (4 months) and the other (half) bonds carry maturity of $2/3t$ (8 months). Further suppose that K is the sum received form the sale of bonds and the firm's average cash holdings equal half the sum ($1/2K$) received from the sale of bonds.

---

13. W.J. Baumol, "Transactions Demand for Cash. An Inventory Theoretic Approach," *Quarterly*

Journal of Economics, November, 1952; and James Tobin, "The Interest Elasticity of the Transactions Demand for Cash," *Review of Economic and Statistics,* August 1956. Our analysis follows Baumol.

Given these assumptions, the firm buys bonds with 2/3*K* ($800) of its income at time *t*=0 and keeps 1/3*K* ($400) in cash, as shown in the figure. At time 1/3*t*, the first half of the bonds purchased ($400) mature which it sells for cash until time 2/3*t*. At time 2/3*t*, the remaining bonds mature which the firm sells for transactions purposes until time t_1. At time t_1, when the year is over, the cash balance is zero and the firm is again ready for fresh receipts in the new year.

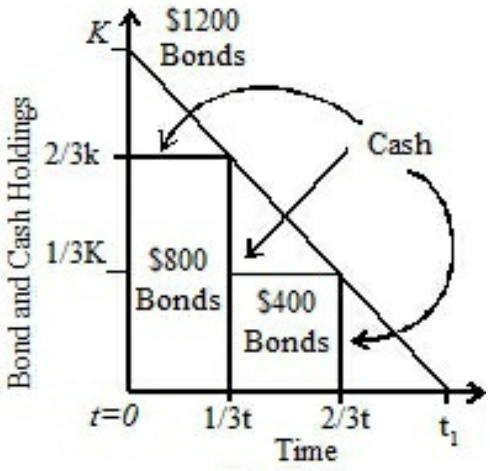


Fig. 7

Now the problem is how to hold assets by a firm, "given that there exist interest-yielding bonds that can be owned as well as cash, and given that there is a fixed cost involved in exchanging bonds for cash."

The solution of this problem requires minimising the cost of holding cash balances over the year. The holding of cash balances consists of interest cost and non-interest costs. Interest cost is in the nature of opportunity cost because when a firm holds cash balances for transactions purposes it forgoes interest income. On the other hand, non-interest costs include such items as brokerage fees, mailing expenses, book-keeping expenses, etc. for converting cash into bonds, and vice versa.

Thus whenever a firm holds money for transactions purposes, it incurs interest costs and brokerage fees (non-interest costs). Let *r* be the rate of interest which is assumed to be constant over the year and *b* the brokerage fee which is also assumed to be fixed. Assume that at the beginning of the year, *Y* is the income of the firm which is equal to the real value of the transactions performed by it, and *K* is the size of each cash withdrawal at intervals over the year when the bonds are sold. Thus *Y*/*K* is the number of withdrawals that occur over the year. The cost on brokerage fees during the year will equal *b*(*Y*/*K*). Since the average cash withdrawal are *K*/2, the

interest cost of holding cash balances is $rK/2$. Then the total cost of making transactions, C, may be written in equation form as:

$$C = r\frac{K}{2} + b\frac{Y}{K} \qquad \ldots(1)$$

The optimal value of K is that which minimises the total inventory cost C. By differentiating C with respect to K, setting the derivative dC/dK equal to zero, and solving for C, we obtain

$$\frac{dC}{dK} = \frac{r}{2} + \frac{-bY}{K^2} = 0$$

or

$$\frac{r}{2} = \frac{bY}{K^2}$$

Multiplying both sides by $2K_2/r$, we obtain

$$K^2 = \frac{2bY}{r}$$

or

$$K = \sqrt{\frac{2bY}{r}} \qquad \ldots(2)$$

Equation (2) shows that if the brokerage fee increases, the number of withdrawals will decrease. In other words, the optimal cash balance will increase because the firm will invest less in bonds. On the other hand, if the rate of interest on bonds rises, the firm will find it profitable to invest in bonds and the optimal cash balance will be lower, and vice versa.

Baumol's analysis points toward another important fact about the behaviour of demand for transactions balances. When a firm or an individual purchases large number of bonds, it is left with small transactions balances and vice versa. But every purchase involves non-interest costs in the form of brokerage fee, mailing, etc. which the purchaser has to pay. He has, therefore, to balance the income to be forgone by making fewer bond purchases against the expenses to be incurred by making large bond purchases. This decision depends upon the rate of interest on bonds. The higher the rate of interest, the larger the expenses which a firm can abosrb in making bond purchases. A more

important factor which determines this decision is the amount of money involved in transactions because brokerage fees of buying and selling bonds are relatively fixed and do not change much in relation to the former. When the money involved in transactions is larger, the smaller will be the brokerage costs. "On a $ 1000 bond purchase, minimum brokerage fees can be costly. On a million dollar transaction they are negligible. Hence, the larger the total amounts involved, the less significant will be the brokerage costs, and the more frequent will be optimal withdrawals." This is because of the operation of economies of scale in cash management or use of money.

It implies that at higher levels of income, the average cost of transactions i.e. brokerage fees are lower. As income increases, the transactions demand for money also increases but by less than the increase in income. If income increases fourfold, optimal transactions balances only double. Since Baumol takes the income elasticity of demand for money to be one-half (1/2), the demand for money will not increase in the same proportion as the increase in income. This is because of the economies of scale that encourage larger investment in bonds when the amount of money involved in transactions is larger due to increase in income.

In this inventory theory of the demand for money, Baumol also emphasises that the demand for money is a demand for real balances. Since the value of average cash holdings over the year is $K/2$, the demand for real balances for transactions purposes becomes

$$\frac{M_D}{P} = \frac{K}{2} = \frac{1}{2}\sqrt{\frac{2bY}{r}}$$

$$M_D = \frac{1}{2}\sqrt{\frac{2bY}{r}}.P \qquad \ldots(3)$$

where M_D is the demand for money and P is the price level.

Equation (3) shows that the demand for real transactions balances "is proportional to the square root of the volume of transactions and inversely proportional to the square root of the rate of interest."

It means that the relationship between changes in the price level and the

transactions demand for money is direct and proportional. The pattern of a firm's purchases remaining unchanged, the optimal cash balances (*Y*) will increase in exactly the same proportion as the price level (*P*). If the price level doubles, the money value of the firm's transactions will also double. When all prices double, brokerage fee (*b*) will also double "so that larger cash balances will become desirable in order to avoid investments and withdrawals and the brokerage costs which they incur." Thus the increase in the money value of transactions and in brokerage fees leads to a rise in the optimal demand for money in exactly the same proportion as the change in the price level. Thus Baumol's analysis of the demand for real balances implies that there is no money illusion in the demand for money for transactions purposes.

Its Superiority over the Classical and Keynesian Approaches

Baumol's inventory theoretic approach to the transactions demand for money is an improvement over the classical and Keynesian approaches.

1. The cash balances quantity theory of money assumed the relationship between the transactions demand and the level of income as linear and proportional. Baumol has shown that this relationship is not accurate. No doubt it is true the transactions demand increases with increase in income but it increases less than proportionately because of the economies of scale in cash management.

2. Baumol's theory also has the merit of demonstrating the interest elasticity of the transactions demand for money as against the Keynesian view that it is interest inelastic.

3. Baumol analyses the transactions demand for real balances thereby emphasising the absence of money illusion.

4. Baumol's inventory theoretic approach is superior to both the classical and Keynesian approaches because it integrates the transactions demand for money with the capital-theory approach by taking assets and their interest and non-interest costs into account.

5. Baumol's theory removes the dichotomy between transactions and

speculative demand for money of the Keynesian approach.

2. Tobin's Portfolio Selection Model: The Risk Aversion Theory of Liquidity Preference

James Tobin in his famous article "Liquidity Preference as Behaviour Towards Risk,"[14] formulated the risk aversion theory of liquidity preference based on portfolio selection. This theory removes two major defects of the Keynesian theory of liquidity preference. *One*, Keynes's liquidity preference function depends on the inelasticity of expectations of future interest rates; and *two*, individuals hold either money or bonds. Tobin has removed both the defects. His theory does not depend on the inelasticity of expectations of future interest rates but proceeds on the assumption that the expected value of capital gains or losses from holding interest-bearing assets is always zero. Moreover, it explains that an individual's portfolio holds both money and bonds rather than only one at a time.

---

14. *The Review of Economic Studies*, February 1958. Reprinted in M.G. Mueller (ed), *Readings* in *Macroeconomics*, 2/e, 1971.

Tobin starts his portfolio selection model of liquidity preference with this presumption that an individual asset holder has a portfolio of money and bonds. Money neither brings any return nor imposes any risk on him. But bonds yield interest and also bring income. However, income from bonds is uncertain because it involves a risk of capital losses or gains. The greater the investment in bonds, the greater is the risk of capital loss from them. An investor can bear this risk if he is compensated by an adequate return from bonds.

If g is the expected capital gain or loss, it is assumed that the investor bases his actions on his estimate of its probability distribution. It is further assumed that this probability distribution has an expected value of zero and is independent of the level of the current rate of interest, r, on bonds.

His portfolio consists of a proportion M of Money and B of bonds where both M and B add up to 1. They do not have any negative values. The return on portfolio R is

$$R = B\ (r+g) \text{ where } O \leq B \leq 1$$

Since *g* is a random variable with expected value zero, the expected return on the portfolio is

$$RE = \mu R = Br.$$

The risk attached to a portfolio is measured by the standard deviation of *R*, that is,σR.

Tobin describes three types of investors. The *first* category is of *risk lovers* who enjoy putting all their wealth into bonds to maximise risk. They accept risk of loss in exchange for the income they accept from bonds. They are like gamblers. The *second* category is of *plungers*. They will either put all their wealth into bonds or will keep it in cash. Thus plungers either go all the way, or not at all.

But the majority of investors belong to the *third* category. They are *risk averters* or *diversifiers*. Risk averters prefer to avoid the risk of loss which is associated with holding bonds rather than money. They are prepared to bear some additional risk only if they expect to receive some additional return on bonds, provided every increase in risk borne brings with it greater increase in returns. They will, therefore, diversify their portfolios, and hold both money and bonds. Although money neither brings any return nor any risk, yet it is the most liquid form of assets which can be used for buying bonds any time.

In order to find out risk averter's preference between risk and expected return, Tobin uses indifference curves having positive slopes indicating that the risk averter demands more expected returns in order to take more risk.

This is illustrated in Figure 8 where the horizontal axis measures risk (σR) and the vertical axis the expected returns ($\sigma\mu R$). The line *Or* is the budget line of the risk averter. It shows the combinations of risk and expected return on the basis of which he arranges his portfolio of wealth consisting of money and bonds. I_1 and I_2 are indifference curves. An indifference

curve shows that he is indifferent between all pairs of expected return and risk that lie on I_1 curve. Points on I_2 curve are preferred to those on I_1 curve. But the risk averter will achieve an equilibrium position between expected return and risk where his budget line is tangent to the indifference curve. It is point *T* on the budget line O_r and I_1 curve.

In the lower portion of the figure, the length of the vertical axis shows the wealth held by the risk averter in his portfolio consisting of money and bonds. The line *OC* shows risk as proportional to the share of the total portfolio held in bonds. Thus point *E* on this line drawn as perpendicular from point T determines the portfolio mix of money and bonds. It is *OP* of bonds shown as *B*, and *PW* of money shown as *M* in the figure.

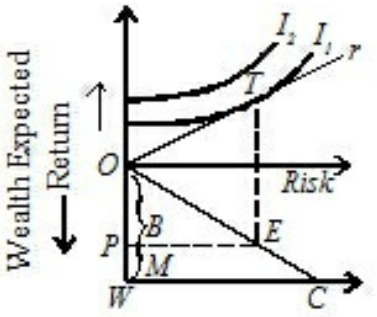


Fig. 8

Thus the risk averter diversifies his total wealth *OW* by putting partly in bonds and partly keeping in cash. That is why he is called a diversifier. He is not prepared to accept more risk unless he can also expect greater expected return. However, the risk averter possesses an intrinsic preference for liquidity which can be only offset by higher interest rates. The higher the interest rate, the lower the demand for money, and the higher the incentive to hold more bonds. On the contrary, the lower the interest rate, the higher the demand for money, and the lower the willingness to hold bonds. This is illustrated in Figure 9.

The slope of the budget line increases with the increase in the interest rate. This is shown by the budget line r_1 rotating upward to r_2 and r_3. Consequently, returns increase in relation to risk with increase in the interest rate, and the budget line touches higher indifferences curves. In Figure 9, budget lines r_1, r_2 and r_3 are tangents to I_1, I_2 and I_3 curves at points T_1, T_2 and T_3 respectively. These points trace out the optimum portfolio curve, *OPC*, in the figure which shows that as the tangency points move upward from left to right, both the expected return and risk increase.

These tangency points also determine the portfolio selection of risk averters as shown in the lower portion of Figure 9. When the rate of interest is r_1, they hold OB_1 bonds and B_1W money. As the rate of interest increases, from r_1 to r_2 and r_3, risk averters hold successively more bonds OB_2 and OB_3 and reduce money to B_2W and B_3W in their portfolios. The figure also shows that as the rate of interest increases by equal increments from r_1, to r_2 to r_3, risk averters hold bonds by decreasing increments, $B_2B_3<B_2B_1<OB_1$. This also means that the demand for money falls by smaller amounts, as the rate of interest increases. This is because the total wealth in the portfolio consists of bonds plus money.

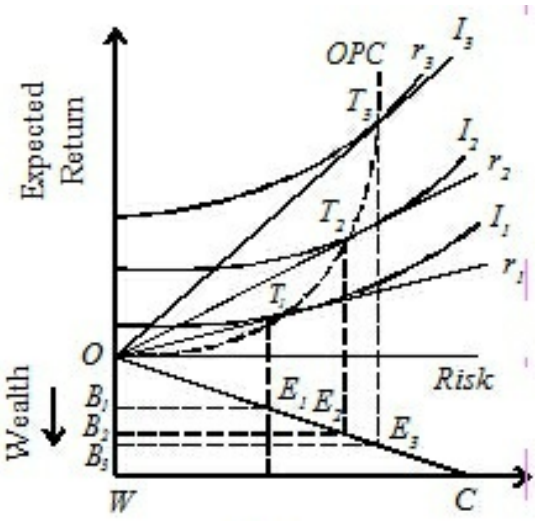


Fig. 9

The demand for money curve can thus be drawn on the basis of figure 9. This is depicted in Figure 10 as the *Ls* curve. The curve shows that when the rate of interest falls from a higher level, there is a smaller increase in the demand for money. For instance, when the interest rate falls from r_{10} to r_8, the demand for money increases by *AB* which is smaller than *OA*. This is because risk averters prefer to hold more bonds than money. But when the rate of interest falls at a lower level from r_4 to r_2, the increase in the demand for money is much larger. It is *CD* in Figure 10. This demand for money curve relates to the speculative demand for money and not to the aggregate demand for money.

Its Superiority over Keynesian Theory

Tobins' risk aversion theory of portfolio selection is superior to the Keynesian liquidity preference theory of speculative demand for money on the following counts:

First, Tobin's theory does not depend on inelasticity of expectations of future interest rates, but proceeds from the assumption that the expected

value of capital gain or loss from holding interest-bearing assets is always zero. In this respect, Tobin regards his theory as a logically more satisfactory foundation for liquidity preference than the Keynesian theory.

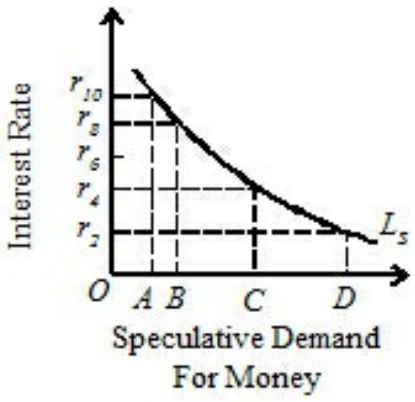


Fig. 10

Second, this theory is superior to Keynes's theory in that it explains that individuals hold diversified portfolios of bonds and money rather than either bonds or money.

Third, like Keynes, Tobin regards the demand for money as closely dependent on interest rates and inversely related to interest rates and his theory provides a basis for liquidity preference.

Fourth, Tobin is more realistic than Keynes in not discussing the perfect elasticity of demand for money (the liquidity trap) at very low rates of interest.

Fifth, according to David Laidler, the real importance of the portfolio theory lies in "not what it tells directly about the aggregate economy, but rather it represents an interesting approach to the problem of relating demand for money to the existence of uncertainty, an approach that probably has scope for considerable development in the future."

EXERCISES

1. What are the motives for holding cash balances according to Keynes ? Give the modifications made by modern economists.
2. Analyse the inventory theory approach to the transactions demand for money. What is its relationship with the rate of interest ?
3. Discuss the portfolio selection approach to the speculative demand for money. How is it superior to Keynes's liquidity preference approach?
4. Is liquidity preference satisfying the speculative demand for money alone which is interest elastic, or other liquidity preferences also ?

Give reasons for your answer.

5. Bring out the relationship between money and interest.

CHAPTER 33

THEORIES OF INTEREST RATE

INTRODUCTION

This chapter discusses some of the important theories of interest rate such as the classical, the loanable funds, the Keynesian, and the modern theory of interest. Besides, certain issues like the natural and market rates of interest are also examined.

Of the theories discussed below, the Keynesian liquidity preference theory that determines the interest rate by the demand for and supply of money is a *stock* theory. It emphasises that the rate of interest is a purely monetary phenomenon. It is a *stock* analysis because it takes the supply of money as given during the short run. and determines the interest rate by liquidity preference or demand for money. On the otter hand, the loanable funds theory is a *flow* theory that determines the interest rate by the demand for and supply of loanable funds. It involves the linking of the interest rate with dissaving, investment and hoarding of funds on the demand side with saving, dishoardings and bank money on the supplyside. These are all flow variables. Hicks and Hansen have reconciled and synthesised these stock and flow theories in a general equilibrium framework and presented a determinate theory of interest rate in terms of the IS-LM formulation.

THE CLASSICAL THEORY OF INTEREST

According to the classical theory, rate of interest is determined by the supply of and demand for capital. The supply of capital is governed by the time preference and the demand for capital by the expected productivity of

capital. Both time preference and productivity of capital depend upon waiting or saving or thrift. The theory is, therefore, also known as the supply and demand theory of saving.

Demand Side. The demand for capital consists of the demand for productive and consumptive purposes. Ignoring the latter, capital is demanded by the investors because it is productive. But the productivity of capital is subject to the law of variable proportions. Additional units of capital are not as productive as the earlier units. A stage comes, when the employment of an additional unit of capital in the business is just worth while and no more. Suppose, an investor invests Rs.1,00,000 in a factory and expects, a yield of 20%. Another instalment of an equal amount would not be as productive as the first one and might bring him 15%. While a third instalment might yield 10%. If he has borrowed the money at 10%, he will not venture to invest more. For the rate of interest is just equal to the marginal productivity of capital to him. It shows that at a higher rate of interest, the demand for capital is low and it is high at a lower rate of interest. Thus the demand for capital is *inversely* related to the rate of interest, and the demand schedule for capital or investment curve slopes downward from left to right There, are, however, certain other factors which govern the demand for capital, such as the growth of population, technical progress, process of rationalization, the standard of living of the community, etc.

Supply Side. The supply of capital depends upon savings, rather upon the will to save and the power to save of the community. Some people save irrespective of the rate of the interest. They would continue to save even if the rate of interest were zero. There are others who save because the current rate of interest is just enough to induce them to save. They would reduce their savings if the rate of interest fell below this level. Still there are the potential savers who would be induced to save if the rate of interest were raised. To the last two categories of savers, saving involves a sacrifice,

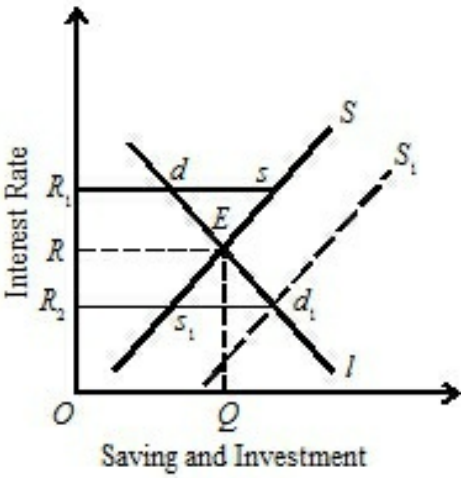


Fig. 1

abstinence or waiting when they forgo present consumption in order to earn interest. The higher the rate of interest, the larger will be the community savings and the more will be the supply of funds. The supply curve of capital or the saving curve thus moves upward to the right

Determination. Assuming the level of income to be given, the rate of interest is determined by the interaction of the demand curve and the supply curve of saving. This is shown in Figure 1 where *I* and *S* curves intersect at *E* which is the equilibrium point when *OQ* quantity of capital is demanded and supplied at *R* rate of interest. If at any time the rate of interest rises above *R,* the demand for investment funds will fall and the supply of savings will increase. Since the supply of savings is more than the demand *(R_1s>R_1 d),* the rate of interest will come down to the equilibrium level *OR.* The opposite will be the case if the rate of interest falls to R_2. The demand for investment funds is greater than the supply of savings *(R_2d_1>R_2s_1)* rate of interest will rise to *R.* The ultimate situation is one of equality between saving and investment brought about by the equilibrium or the *natural* rate of interest. If at any time people become thrifty and save more than *OQ,* the rate of interest would fall below *R* because the demand for capital remains the same. This is shown by the downward shift of the saving curve to S_1, where it intersects the *I* curve at d_1 and the rate of interest falls to R_2 At the lower rate of interest, people will save less but the demand for investment funds will increase which will tend to raise the rate of interest to the equilibrium level *R.*

Its Criticisms

The 'pure' or the *real* theory of interest of the classicals, as enunciated by Marshall and Pigou, has been severely criticised by Keynes.

(1) Income not Constant but Variable. One of the serious defects of the classical theory is that it assumes the level of income to be given, and regards interest as an equilibrating mechanism between the demand for investible funds and the supply of funds through savings. According to Keynes, income is a variable and not a constant and the equality between saving and investment is brought about by changes in income and not by variations in the rate of interest.

(2) Saving-Investment Schedules not Independent. In this theory the two determinants of interest rate, the demand and supply curves of saving are treated as independent of one another. It means that if there is change in demand, the demand curve for savings can shift up or below the *I* curve without causing a change in the supply curve. But according to Keynes, the two curves are not independent of one another. If, for instance, an invention shifts the investment curve upward, income will rise and it will lead to higher savings and thus shift the supply curve too. Similarly, a shift in the supply curve will bring a change in the demand curve.

(3) Neglects the Effects of Investment on Income. The classical theory neglects the effect of investment on the level of income. A rise in the rate of interest, for instance, will bring a decline in investment by making it less profitable. This will mean decline in output, employment and income. The latter will, in turn, lead to reduced savings, a fact contrary to the classical assertion that saving is a direct function of the rate of interest. On the other hand, a low rate of interest encourages investment activity, increases output, employment, income and savings. But Keynes does not believe that investment depends on the rate of interest. It depends on the marginal efficiency of capital. Even if the rate of interest were to fall to zero, Keynes argues, investment will not take place if business expectations for profits are at a low level, as is the case in depression.

(4) Indeterminate Theory. Since savings depend upon the level of income, it is not possible to know the rate of interest unless the income level is known beforehand. And the income level itself cannot be known without already knowing the rate of interest. A lower rate of interest will increase investment, output, employment, income and savings. So, for each income level a separate saving curve will have to be drawn. This is all circular reasoning and offers no solution to the problem of interest. That is why Keynes characterised the classical theory of interest as indeterminate.

(5) Neglects other Sources of Savings. The propounders of this theory include savings out of current income in the supply schedule of savings which makes it inadequate. Considering the supply of capital to be interst-elastic, people might lend their past savings with the rise in the rate of

interest and so increase the supply of capital. Similarly, bank credit is an important source of the supply of capital. Banks lend more during periods of slow business activity. The classical theory remains incomplete when it neglects these factors in the supply schedule of capital.

(6) Unrealistic Assumption of Full Employment. The classical theory is based on the unrealistic assumption of full employment. In a fully employed economy interest as a reward for saving, waiting or abstinence is necessary to induce people to save. But according to Keynes, underemployment and not full employment is the rule and where resources are unemployed, interest is not essentially an inducement to savings.

(7} Neglects Monetary Factors. The classical theory is a pure or real theory of interest which takes into consideration the real factors like the time preference and the marginal productivity of capital. It completely neglects the influence of monetary factors on the determination of the rate of interest. The classical economists regarded money as a veil, a medium of exchange over goods and services. They failed to consider it as a store of value. Keynes, on the other hand, laid emphasis on explaining the determination of the rate of interest as a monetary phenomenon.

(8) No Automatic Equality between Equilibrium and Market Rates of Interest.

According to the classical view, the market and the equilibrium (natural) rates of interest are always equal. Any discrepancy between the two is only a temporary phenomenon which would disappear in the long run. Keynes, however, does not regard the discrepancy between the two as accidental and temporary. It can be due to the contraction or expansion of bank credit. An expansion of bank credit by increasing the supply of loanable funds brings about a fall in the market rate of interest below the equilibrium rate and *vice versa.* Thus there is no automatic mechanism for the equality of the market and equilibrium interest rate.

(9) Difference over the Definition of Interest. Keynes differs with the classical economists even over the definition and determination of the rate of interest. According to him, it is the reward of not hoarding but the

reward of parting with liquidity for a specified period. It is the 'price' which equilibrates demand for money with the available quantity of money. He does not agree that it is determined by the demand for and supply of capital. Thus, Keynes dismisses the classical theory of interest as absolutely wrong and inadequate.

THE LOANABLE FUNDS THEORY OF INTEREST

The neo-classical or the loanable funds theory explains the determination of interest in terms of demand and supply of loanable funds or credit.

According to this theory, the rate of interest is the price of credit, which is determined by the demand and supply for loanable funds. In the words of Prof. Lerner, it is the price which equates the supply of 'credit', or saving *plus* the net increase in the amount of money in a period, to the demand for 'credit', or investment *plus* net hoarding "in the period." Let us analyse the forces behind the demand and supply of loanable funds.

Demand for Loanable Funds. The demand for loanable funds has primarily three sources *i.e.,* government, businessmen and consumers who need them for purposes of investment, hoarding and consumption. The government borrows funds for constructing public works or for war preparations. The businessmen borrow for the purchase of capital goods and for starting investment projects. Such borrowings are interest elastic and depend mostly on the expected rate of profit as compared with the rate of interest. The demand for loanable funds on the part of consumers is for the purchase of durable consumer goods like scooters, houses, etc. Individual borrowings are also interest elastic. The tendency to borrow is more at a lower rate of interest than at a higher rate in order to enjoy their consumption soon. Since this demand for funds is mostly met out of past savings or through *dis-saving,* it is represented by the curve *Ds* in Figure 2. The demand curve for *investment funds,* both for the government and the businessmen is shown as curve *I.* It slopes downward showing that less funds are borrowed at a higher rate and more at a lower rate of interest. *Lastly,* funds are demanded for the purpose of *hoarding* them in liquid form or as idle cash. They are also interest elastic and are shown by the curve *H.* The lateral summation of these curves *H, Ds* and *I* gives us

the aggregate demand curve for loanable funds ΣD.

Supply of Loanable Funds. The supply of loanable funds comes from savings, dishoardings and bank credit. Private, individual and corporate savings are the main source of savings. Though personal savings depend upon the income level, yet taking the level of income as given they are regarded as interest elastic. The higher the rate of interest, the greater will be the inducement to save and *vice versa*. Corporate savings are the undistributed profits of a firm which also depend on the current rate of interest to some extent. If the interest rate is high, it will act as a deterrent to borrowing and thus encourage savings. *Savings* are indicated as curve *S* in the Figure 2. The second source is the volume of funds coming out of hoards or being added to them. *Dishoarding* may represent not only purchase of old assets or securities from others out of idle cash balances of one's own funds for net investment or for consumption in purchases in excess of net disposable income. Such funds are directly related to the rate of interest. The higher the interest rate the larger the funds that will be coming out of hoards and *vice* versa. These funds are represented by the curve D_H. Lastly, there is the *bank credit* as an important source of the supply of loanable funds. Bank credit or money is also interest elastic to some extent. More funds are lent at a higher than at a lower rate of interest. Bank credit is shown as the curve *M*. If these curves D_H, *M* and *S* are laterally added up, we have the aggregate supply curve *ΣS* of loanable funds.

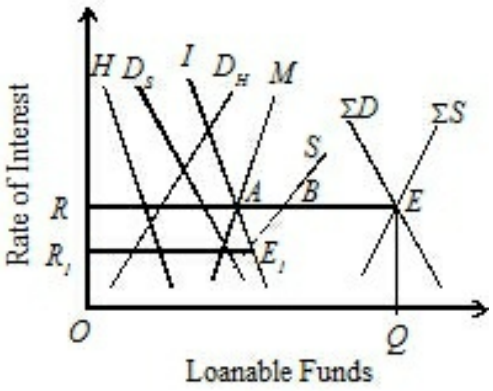


Fig. 2

The total demand curve for loanable funds *ΣD* and the total supply curve of loanable funds *ΣS* intersect at *E* and give *OR* rate of interest. At this rate *OQ* amount of funds are borrowed and lent.

Its Criticisms

According to Prof. Robertson, the loanable funds theory is a "commonsense explanation" of the determination of the rate of interest. But this theory is also not free from certain defects.

(1) Equilibrium Rate reflects Unstable Equilibrium. The demand and supply schedules for loanable funds determine the equilibrium rate of interest *OR* which does not equate each component on the supply side with the corresponding component on the demand side. Thus the equilibrium rate *OR* reflects unstable equilibrium. For stable equilibrium, it is essential that *ex ante* (planned) investment must equal *ex-ante* savings at the equilibrium rate *OR*. In the figure, *ex-ante* savings *S* exceed *ex-ante* investment *I* by *AB.* They are equal at point E_1 but at a lower rate OR_1 which is the natural rate of interest.

(2) Indeterminate Theory. Prof. Hansen asserts that the loanable funds theory like the classical and the Keynesian theories of interest is indeterminate. The supply curve of loanable funds is composed of savings, dishoardings and bank money. But since savings vary with past income and the new money and activated balances with the current income, it follows that the total supply curve of loanable funds also varies with income. Thus the loanable funds theory is indeterminate unless the income level is already known.[1]

(3) Cash Balances not Elastic. The loanable funds theory states that the supply of loanable funds can be increased by releasing cash balances of savings and decreased by absorbing cash balances into savings. This implies that the cash balances are fairly elastic. But this does not seem to be a correct view because the total cash balances available with the community are fixed and equal the total supply of money at any time. Whenever there are variations in the cash balances, they are in fact in the velocity of circulation of money rather than in the amount of cash balances with the community.

(4) Savings not Interest Elastic. The theory over-emphasises the influence of the rate of interest on savings. It regards savings as interest elastic. Generally speaking, people save not to earn rate of interest but to satisfy precautionary motive. So savings are interest *inelastic.*

(5) Not correct to combine Real and Monetary Factors. The loanable funds theory has been criticised for combining monetary factors with real factors. It is not correct to combine real factors like saving and investment

with monetary factors like bank credit and dishoarding without bringing in changes in the level of income. This makes the theory unrealistic.

Its Superiority over the Classical Theory

Despite these weaknesses, the loanable funds theory is better and more realistic than the classical theory on a number of counts.

1. The classical theory is a real theory of interest and neglects monetary influences on interest. With the inclusion of real as well as monetary factors, the loanable funds theory becomes superior to the classical theory.

2. The classicists neglect the role of bank credit as a constituent of money supply influencing the rate of interest which is an important factor in the loanable funds theory.

3. The classicists also do not consider the role of hoarding. By including the desire to hoard money in the demand for loanable funds, the loanable funds theory becomes more realistic and brings us nearer to Keynes's liquidity preference theory.

4. To the classicists money is merely a 'veil,' a passive factor influencing the rate of interest. The loanable funds theory is superior because it regards money as an active factor in the determination of the interest rate.

---

1. A. H. Hansen*A Guide to Keynes* p. 141.

KEYNES'S LIQUIDITY PREFERENCE THEORY OF INTEREST

Keynes defines the rate of interest as the reward of not hoarding but the reward for parting with liquidity for the specified period. It "is not the 'price' which brings into equilibrium the demand for resources to invest with the readiness to abstain from consumption. It is the 'price' which equilibrates the desire to hold wealth in the form of cash with the available quantity of cash."[2] In other words, the rate of interest, in the Keynesian sense, is determined by the demand for and the supply of money. This theory is, therefore, charaterised as the monetary theory of interest, as

distinct from the real theory of the classicals.

Supply of Money. Of the two determinants of the rate of interest, the supply of money refers to the total quantity of money in the country for all purposes at any time. Though the supply of money is a function of the rate of interest to a degree, yet it is considered to be fixed by the monetary authorities, that is, the supply curve of money is taken as perfectly inelastic.

Demand for Money. For the second determinant, the demand for money, Keynes coined a new term "liquidity preference" by which his theory of interest is commonly known. Liquidity preference is the desire to hold cash. The money in cash "lulls our disquietude" and the rate of interest which is demanded in exchange for it is a "measure of the degree of our disquietude." The rate of interest, in Keynes words, is the "premium which has to be offered to induce people to hold the wealth in some form other than hoarded money." The higher the liquidity preference,the higher will be the rate of interest that will have to be paid to the holders of cash to induce them to part with their liquid assets. The lower the liquidity preference, the lower will be the rate of interest that will be paid to the cash-holders. According to Keynes, there are three motives behind the desire of the people to hold liquid cash: (1) the transaction motive, (2) the precautionary motive, and (3) the speculative motive.

Transactions Motive. The transactions motive relates to "the need of cash for the current transactions of personal and business exchanges." It is further divided into the income and business motives. The *income motive* is meant to bridge the interval between the receipt of income and its disbursement," and similarly, the *business motive* as "the interval between the time of incurring business costs and that of the receipt of the sale proceeds." If the time between the incurring of expenditure and receipt of income is small, less cash will be held by the people for current transactions, and vice versa. There will however be changes in the transactions demand for money depending upon the expectations of the income, of recipients and businessmen. They depend upon the level of income, employment and prices, the business turnover, the normal period between the receipt and disbursement of income, the amount of salary or

income, and on the possibility of getting a loan.

---

2. J.M Keynes, *The General Theory of Employment, Interest and Money, p.* 167.

Precautionary Motive. The precautionary motive relates to "the desire to provide for contingencies requiring sudden expenditures and for unforeseen opportunities of advantageous purchases." Both individuals and businessmen keep cash in reserve to meet unexpected needs. Individuals hold some cash to provide for illness, accidents, unemployment and other unforeseen contingencies. Similarly businessmen keep cash in reserve to tide over unfavourable conditions or to gain from unexpected deals. Money held under the precautionary motive is rather like water kept in reserve in a water tank. The precautionary demand for money depends upon the level of income, and business activity, opportunities for unexpected profitable deals, availability of cash, the cost of holding liquid assets in bank reserves, etc.

Keynes holds that the transactions and precautionary motives are relatively *interest inelastic,* but are highly *income elastic.* The amount of money held under these two motives (M_1) is a function (L_1) of the level of income (Y) and is expressed as $M_1=L_1 (Y)$.

Speculative Motive. Money held under the speculative motive is for "securing profit from knowing better than the market what the future will bring forth." Individuals and businessmen have funds, after keeping enough for transactions and precautionary purposes and like to gain by investing in bonds. Money held for speculative purposes is a liquid store of value which can be invested at an opportune moment in interest-bearing bonds or securities. Bond prices and the rate of interest are inversely related to each other. Low bond prices are indicative of high interest rates, and high bond prices reflect low interest rates. A bond carries a fixed rate of interest. For instance, if a bond of the value of Rs 100 carries 4% interest and the market rate of interest rises to 8%, the value of this bond falls to Rs 50 in the market. If the market rate of interest falls to 2%, the value of the bond will rise to Rs 200 in the market. Thus individuals and businessmen can gain by buying bonds worth Rs 100 each at the market rate of Rs 50 each when the rate of interest is high (8%), and sell them

again when they are dearer (Rs 200 each) when the rate of interest falls (to 2%).,

According to Keynes, it is expectations about changes in bond prices or in the current market rate of interest that determine the speculative demand for money. The speculative demand for money is a decreasing function of the rate of interest. The higher the rate of interest, the lower the speculative demand for money, and lower the rate of interest, the higher the speculative demand for money. Algebraically, Keynes expressed the speculative demand for money as $M_2=L_2(r)$ where L_2 is the speculative demand for money and r is the rate of interest. Geometrically, it is a smooth curve which slopes downward from left to right, as shown in Figure 3 .

But at a very low rate if interest, such as 2% the speculative demand for money becomes perfectly elastic. This portion of the curve is known as the *liquidity trap.* At a very low rate of interest, people prefer to keep money in cash rather than invest in bonds because purchasing bonds will mean a definite loss.

Total Demand of Money

If the total liquid money is denoted by *M*, the transactions plus precautionary motive by M_1 and the speculative motive for holding by M_2, then$M=M_1+M_2$. Since $M_1 = L_1(Y)$ and$M_2=L_2(r)$, the total liquidity preference function is expressed as $M=L(Y,r)$. M_1 is circulating or active money and M_2 is idle, or passive money. Though M_1 is a function of income and M_2 of the rate of interest, yet they cannot be held in water-tight compartments. Even M_1 is interest elastic at high interest rates. If there is increased demand for M_1 it can be met by transferring funds from idle balances, M_2. Prof. Haberler distinguishes between liquidity preference in the narrower sense and liquidity preference in the wider sense. The former is the demand for idle balances, M_2 and is called "liquidity preference proper", and the latter is the total liquidity demand *(M or* M_1+M_2). Given the level of income at high interest rates, liquidity preference refers to the total demand for money (M_1+M_2) and at low

interest rates the demand for speculative motive *(M_2)*alone.

Determination of the Rate of Interest. We have analysed the factors behind the supply and demand for money. The rate of interest, like the price of any product or service, is determined at a level where the demand for money equals the available supply of money. In Figure 3, the vertical line Q_1M represents the supply of money and *L* is the total demand for money curve. Both intersect at *E* where the *equilibrium rate* of interest *R* is established. If there is any deviation from this equilibrium position, an adjustment will take place via the rate of interest, and the equilibrium level *E* is re-established. E_1 is the position of unstable equilibrium where the supply of money *OM* is greater than the demand for moneyOM_1.Consequently, the rate of interest will start declining from R_1, till the equilibrium rate of interest *R* is reached. Similarly at R_2 level of interest rate, the demand for money OM_2 is greater than the supply of money *OM.* As a result, the rate of interest R_2 will start rising till it reaches the equilibrium rate *R.*

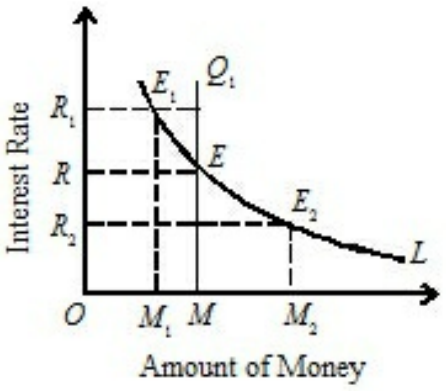


Fig. 3

If the *supply of money* is increased by the monetary authorities, but the liquidity preference curve *L* remains the same, the rate of interest will fall. This is illustrated in Figure 4. Given the *L* curve, the supply of money curve being *QM,* the rate of interest is R_5. With the increase in the supply of money from *QM* to O_1M_1 and Q_2M_2, the rate of interest falls from R_3 to R_2. But any further increase in the supply of money has no effect on the rate of interest because the liquidity preference curve *L* is perfectly elastic at R_2 rate of interest. So when the

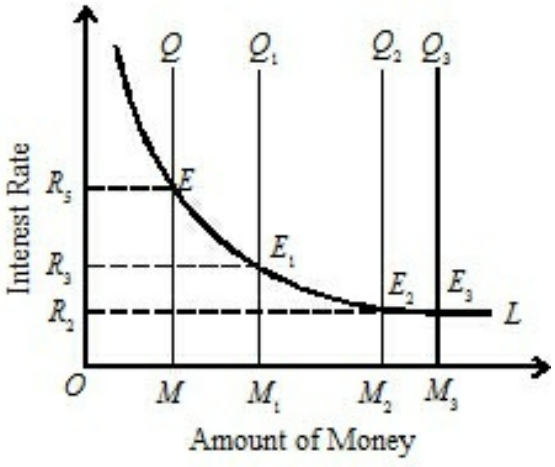


Fig. 4

supply of money increases to Q_3M_3, the rate of interest remains stationary at R_2 corresponding to the equilibrium point E_3. This is Keynes "Liquidity Trap".

If the demand for money increases and the liquidity preference carve shifts upward, given the supply of money, the rate of interest rises. This is shown in Figure 5. Given the supply of money curve *QM* when the *L* curve shifts upward, the new equilibrium point is E_1 which determines *R*, $_6$interest rate. This rate of interest is higher than R_4 interest rate at the equilibrium point *E*. If with the increase in the liquidity preference, the supply of money also increases in the same proportion to Q_1M_1 there will be no change in the rate of interest R_4, except that the new equilibrium point is E_2. Thus the theory explains that the rate of interest is determined at a point where the liquidity preference curve equals the supply of money curve.

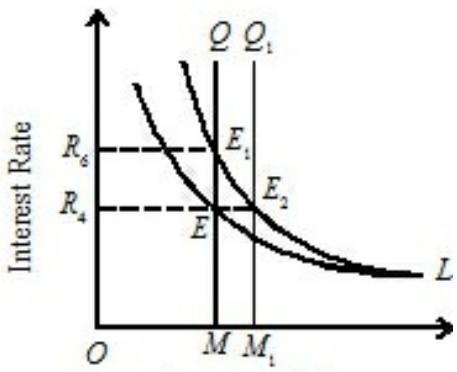


Fig. 5

Its Criticisms

The Keynesian theory of interest has been severely criticised by Hansen, Roberston, Knight, Hazlitt, Hutt and others. It has been variously characterized as "a college bursar's theory", "at best an inadequate and at worst a misleading account, and "preclassical, mercantilist, and man-in-the-street economics."

(1) College Bursar's Theory. One of the fallacies of the Keynesian analysis is that the demand for money is mainly associated with the liquidity preference for the speculative motive to which the rate of interest is brought directly into touch. The theory presumes the existence of a number of wealthy and shrewd persons who will hold more money by selling bonds when the rate of interest falls and hold less cash but more bonds in the case of rise in the interest rate. But Robertson does not regard bonds as the only alternative to money, as a use for resources both by the

individual and the entrepreneur. And a theory of money which insists on working everything through the bond market is a *College Bursar's theory* and seems to be lacking in realism and comprehensiveness.

(2) Inadequate and Misleading Theory. The theory tells us nothing as to what determines the normal rate and therefore, given the degree of divergence, the actual rate of interest. The actual rate of interest cannot be described consistently as measuring the cost of the uncertainty of risk involved in holding bonds rather than money. If there were no uncertainty at all, the actual rate would simply stand at a certain definite level below the normal rate. It is for these reasons that Robertson regards the liquidity preference theory "is at best an inadequate and at worst a misleading account."

(3) Falls into Methodological Fallacy. Just as the rate of interest is explained as the price for parting with liquidity, similarly the price of eggs or any other commodity can be explained by the relative preference for them. But there is a vital difference. A change in the quantity of money would tend to change the price of the good in the same proportion, but not the price of bonds beyond a temporary disturbance. In fact, there is no functional relationship between the price level and any rate of interest while discussing the influence of speculation on the interest rate on loans. As a result, no monetary change has any direct or permanent effect on the rate of interest. Thus the Keynesian theory falls into *"methodological fallacy"* by assuming a definite functional relationship between the quantity of money and the rate of interest.

(4) Money as a Store of Wealth is Barren. Keynes holds that only money held for speculative purposes is fruitful for it brings interest as a reward, while money as a store of wealth is barren. This is a mistaken notion. As pointed out by W.H. Hutt, money "is as productive as all other assets, and productive in exactly the same sense. The demand for money assets is a demand for productive resources.

(5) Inconsistent Theory. Knight criticised Keynes's theory in view of the facts which are directly contrary to what the theory calls for. Hazlitt also discusses this point without giving credit to Knight. According to the Keynesian theory, the rate of interest should be the highest at the bottom

of the depression because the liquidity preference is the strongest at that time due to falling prices. A large reward will have to be paid to induce wealth-holders to part with their cash. But the facts are just the opposite. Short-term interest rates are the lowest in a depression because investment opportunities are temporarily closed and the lenders have no outlets to part with their cash. Contrariwise, the Keynesian short-term interest rates should be lowest during the peak of the boom because people would be investing their money rather than keep it in cash. The liquidity preference being the lowest, a very small reward would be required to partake with it. But in reality, the rate of interest is the highest at the peak of a boom. Thus the theory is inconsistent with facts.

(6) Saving Essential for Liquidity. Keynes regards the rate of interest as the reward for parting with liquidity and not a return merely for saving or waiting as such. Saving is necessary for obtaining funds to be invested at interest. In Viner's words "Without saving there can be no liquidity to surrender. The rate of interest is the return for saving without liquidity."

(7) Liquidity not Essential for Interest Rate. Even the term liquidity preference is neither helpful nor necessary in explaining the nature of interest. It is more confusing and less illuminating. It is not only vague but is also self-contradictory. For, as pointed out by Hazlitt, "If a man is holding his funds in the form of time deposits or short-term treasury bills, he is being paid interest on them , therefore he is getting interest and "liquidity" too. What becomes, then, of Keynes's theory that interest is the "reward" for parting with liquidity ?"

(8) Wrong Notion of Liquidity Trap. Keynes's notion of the "Liquidity Trap" is also wrong. In reality, the liquidity preference schedule may be perfectly *inelastic* rather than elastic at a low rate of interest. We know that during a depression all expectations are extremely pessimistic. It is, therefore, not correct to argue that expectations with regard to the rate of interest will be that it will go up.

(9) Ignores Real Factors. The greatest fallacy in Keynes's analysis is that he ignores the influence of real factors in determining the interest rate. He regards the interest rate as a purely monetary phenomenon and thus merely returns "to the pre-classical assumption of the mercantilists and to

what has always been the assumption of the man in the street." According to Hazlitt, "Keynes made no new contribution. He merely muddled shallow waters *and* the kind of *interest* theory represented by *him* is pre-classical, mercantilistic and man in the street economics."

(10) Indeterminate Theory. The Keynesian theory, like the classical theory of interest, is *indeterminate.* Keynes asserts that the liquidity preference and the quantity of money determine the rate of interest. But this is not correct because a new liquidity preference curve will have to be drawn at each level of income. Therefore, unless the income level is already known, the demand and supply curves of money cannot tell us what the rate of interest will be. Thus, according to Prof. Hansen, "Keynes's criticism of the classical theory applies equally to his own theory."

(11) Incomplete Theory. Hicks, Somers, Lerner, Hansen and others opine that the rate of interest alongwith the level of income is determined by four factors: (i) the investment demand function *(MEC), (ii)* The saving function (or the consumption function), *(iii)* the liquidity preference function, and (iv) the quantity of money function. Though all these elements are found in the Keynesian analysis, yet Keynes does not bring them in his interest rate theory. He takes only the last two elements and ignores the first two, thus Keynes fails to provide an integrated and determinate theory of interest.

(12) Confusion regarding Relation between Interest Rate and Quantity of Money. There is confusion in Keynes's analysis about the relation between rate of interest and amount of money. On the one hand, he says that the demand for money is inversely dependent on the rate of interest, and on the other, that the equilibrium rate of interest is inversely dependent upon the amount of money. Throughout his analysis, Keynes does not make any distinction between the two propositions and often uses them in an identical manner This is a fundamental error in Keynes's analysis for the former relation holds true for an individual and the latter for the market.

Conclusion. The Keynesian theory of interest is not only indeterminate, but is also an inadequate explanation of the determination of the rate of

interest. It treats the interest rate as a purely monetary phenomenon and by neglecting the real factors makes the theory narrow and unrealistic.

Its Superiority over the Loanable Funds Theory

Despite these criticisms of the Keynesian theory, it is considered superior to the loanable funds theory on the following counts:

(1) The Keynesian theory is a *stock* analysis. It is a statement about stocks or quantities of money at a point of time, while the loanable funds theory is a statement about certain *flows* or quantities of money per time period. It is this treatment of stocks and flows of money that leads to a fixed supply of money in the Keynesian analysis, and a variable supply of money in the loanable funds theory. Since the quantity of money is fixed at a point of time, economists prefer the stock approach to the rate of interest. Thus the liquidity preference theory becomes superior to the loanable funds theory.

(2) The liquidity preference theory is more realistic than the loanable funds theory because it is more akin to the behaviour of interest rate in the business world. It clearly spells out the various motives for holding money, and the relation between business, expectations and the rate of interest.

(3) All the variables in the loanable funds theory like saving, investment, hoarding, dishoarding are in terms of partial equilibrium analysis while the demand for and supply of money in the Keynesian system have been treated as a part of the general determinate system. So the loanable funds theory fails to fit in a general determinate system.

(4} In the loanable funds theory, saving and investment are redundant and the rate of interest can be determined by $H+D_S$ and $M+D_H$ whereas in the Keynesian theory savings are interest inelastic and investment funds depend upon the money supply. The Keynesian theory is thus superior to the loanable funds theory of interest.

Despite these apparent differences and the superiority of one over the other, there have been three attempts to prove that the two theories of the

interest rate are two ways of looking at the same mechanism. "Like the frugal horse-man who brought only one spur because when one side of the horse went ahead, the other would follow, we could get along with only one of the approaches to the interest rate," as pointed out by a supporter of these views of Hicks, Lerner, Fellner and Somers. Dr. Klein who has critically analysed these attempts at reconciliation of the two theories comes to the conclusion that if only the two theories are explained in terms of stock analysis that they will come to the same thing. Since the loanable funds version is a flow analysis of the interest rate determination, it cannot be reconciled with the liquidity preference theory.[3]

INDETERMINACY OF THE CLASSICAL, THE LOANABLE FUNDS AND THE KEYNESIAN THEORIES OF INTEREST

Keynes criticised the classical theory of interest for being *indeterminate* because it failed to relate the rate of interest with the income level. To Hansen, "Keynes's criticism of the classical theory applies equally to his own theory and to the loanable funds theory." We illustrate the indeterminate nature of these theories.

In the *classical* formulation, since savings depend upon the level of income, it is not possible to know the rate of interest unless the income level is known beforehand. And the income level cannot be known without already knowing the rate of interest. A lower rate of interest will increase investment, output, employment, income and savings. So, for each income level a separate supply curve will have to be drawn.

The same reasoning applies to the *loanable funds* formulation of the rate of interest. The supply schedule of loanable funds is composed of savings, dishoarding and bank money supply. Since savings vary with past income and the new money and activated balances with the current income, it follows that the total supply schedule of loanable funds also varies with income. Thus this theory is *indeterminate* unless the income level is already known. The indeterminate nature of both the theories is explained with the help of Figure 6. All that the classical formulation gives us is a *family* of saving schedules, and the loanable funds formulation a *family* of investment schedules at various income levels relating them to different

interest rates. We take saving and investment (or demand and supply of loanable funds) on the X-axis and the rate of interest on the vertical axis. The savings schedules (for both the formulations) are shown as S_1Y_1 and S_2Y_2 and *II* is the investment demand schedule. When income is Y_2 , given the investment demand schedule and *II*, and the savings schedule S_2Y_2,savings equal investment at interest rate *r*. Similarly when income is Y_1, the savings schedule S_1Y_1 equals the investment demand schedule *II* at interest r_t. These equilibrium positions tell us about the various levels of income associated with different rates of interest, but not about the rate of interest itself. They show that interest is a function of savings, investment and the level of income. Unless the level of income is known, it is not possible to determine the rate of interest. Hence the classical and the loanable funds theories of interest rate are *indeterminate.*

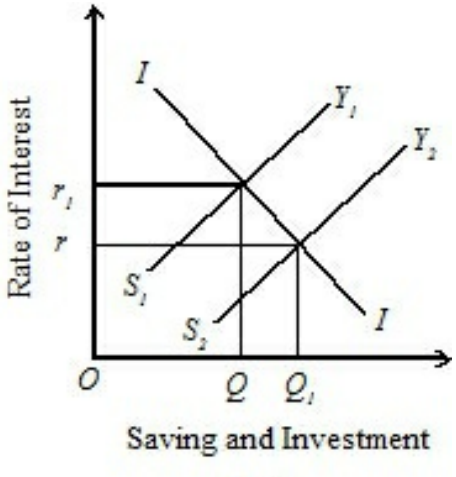


Fig. 6

---

3. L..R. Klein, *The Keynesian Revolution, pp.* 118-20.

The *Keynesian* theory of interest rate is also indeterminate because the liquidity preference schedule is not related to the level of income. Unless the income level is known beforehand, the demand and supply curves of money cannot tell us what the rate of interest will be. All that the Keynesian formulation provides us is a *family* of liquidity preference schedules at various income levels relating them to different interest rates. In Figure 7, a *family* of liquidity preference schedules *LY*, L_tY_1 and and L_2Y_2 is drawn at various income levels. A perfectly inelastic money supply curve *QM* is drawn on the assumption that the supply of money is fixed by the monetary authority. If Y is the income level, the liquidity preference schedule *LY*

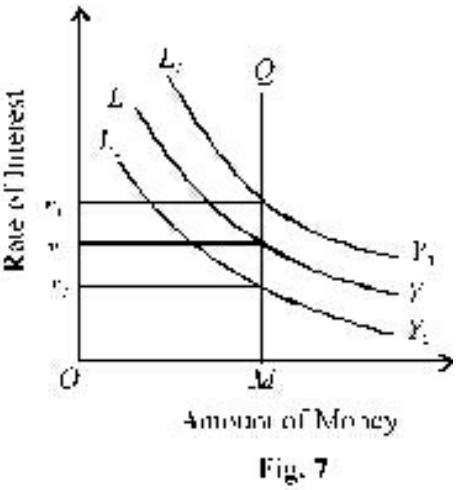


Fig. 7

equals the money supply schedule *QM* at interest rate *r*. If the income level rises to Y_1, the liquidity preference schedule also shifts upward to L_1M_t and equals the *QM* curve at the interest rate *r*. If income falls to Y_2 level, the liquidity preference curve shifts to L_2Y_2 and equals *QM* curve, at the interest rate r_2. Thus the Kenesian theory simply relates different income levels to various interest rates, but does not show what the rate of interest will be. Hence it is an *indeterminate* theory.

MODERN THEORY OF INTEREST

We have seen above that no single theory of interest rate is adequate and determinate. An adequate theory to be determinate must take into consideration both the real and monetary factors that influence the interest rate. Hicks has utilized the Keynesian tools in a method of presentation which shows that productivity, thrift, liquidity preference and money supply are all necessary elements in a comprehensive and determinate interest theory. According to Hansen, "An equilibrium condition is reached when the desired volume of cash balances equals the quantity of money, when the marginal efficiency of capital is equal to the rate of interest and finally, when the volume of investment is equal to the normal or desired volume of saving. And these factors are inter-related."[4] Thus in the modern theory of interest rate, saving, investment, liquidity preference and the quantity of money are integrated at various levels of income for a synthesis of the loanable funds theory with the liquidity preference theory.

The four variables of the two formulations have been combined to construct two new curves, the *IS* curve representing *flow* variable of the loanable funds formulation (or the real factors of the classical theory) and the *LM* curve representing the *stock* variables of liquidity preference formulation.The equilibrium between *IS* and *LM* curves provides a *determinate* solution.

The IS Curve

The *IS* curve has been derived from the loanable funds formulation. It is a curve which explains the relationship between a family of saving

schedules and investment schedules. In other wards, this curve shows the equality of saving and investment at various combinations of the levels of income and the rates of interest. In Figure 8 (A), the saving curve *S* in relation to income is drawn in a fixed position, since the influence of interest on saving is assumed to be negligible. The saving curve shows that saving increases as income increases, viz., saving is an *increasing* function of income. Investment, on the other hand, depends on the rate of interest and the level of income. Given a level of interest rates, the level of investment rises with the level of income. At a 5 per cent rate of interest, the investment curve is I_2 . If the rate of interest is reduced to 4 per cent, the investment curve will shift upward to I_3. The rate of investment will have to be raised to reduce the marginal efficiency of capital to equality with the lower rate of interest. Thus the investment curve I_3 shows more investment at every level of income. Similarly when the interest rate is raised to 6 per cent, the investment curve will shift downward to I_1. The reduction in the rate of investment is essential to raise the marginal efficiency of capital to equality with the higher interest rate. In Figure 8 (B), just below Figure 8 (A), we derive the *IS* curve by marking the level of income at various interest rates. Each point on this *IS* curve represents a level of income at which saving equals investment at various interest rates. The rate of interest is represented on the vertical axis and the level of income on the horizontal axis. If the rate of interest is 6 per cent, the *S* curve intersects the I curve at E_t which determines OY_1 income. From this income level which equals Rs 100 crores we draw a dashed line downward to intersect the extended line from 6 per cent at point A. At interest rate 5 per cent, the *S* curve intersects the I_2 curve at E_2 so as to determine OY_2 income (Rs 200 crores). In the lower Figure 8 (B),

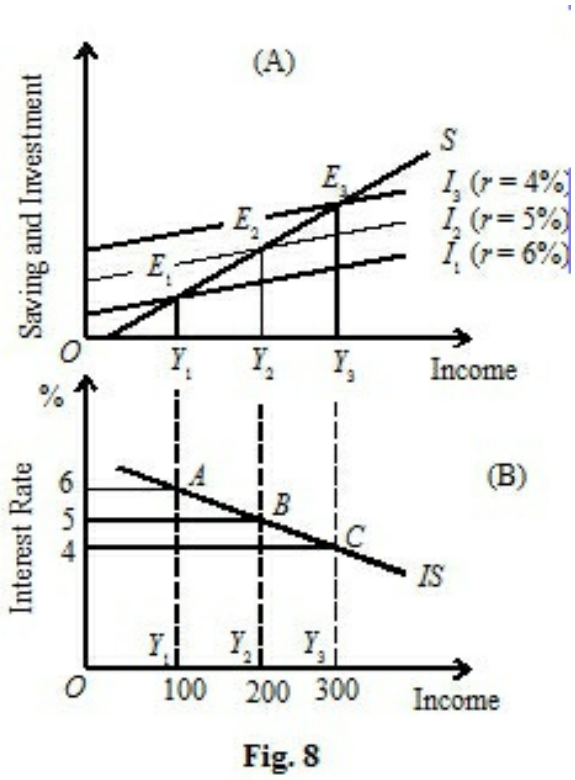


Fig. 8

the point *B* corresponds to 5 per cent interest rate and Rs 200 crores income level. Similarly, the point *C* corresponds to the equilibrium of *S* and I_3 at 4 per cent interest rate. By connecting these points *A, B* and *C* with a line, we get the *IS* curve. The *IS* curve slopes *downward from* left to right because as the interest rate falls, investment increases and so does income.

---

4. A.H. Hansen, *Monetary Theory and Fiscal Policy*, p. 71

The LM Curve

The *LM* curve shows all combinations of interest rates and levels of income at which the demand for and supply of money are equal. The *LM* curve is derived from the Keynesian formulation of liquidity preference schedules of supply of money. A family of liquidity preference curves L_tY_1,L_2Y_2 and L_3Y_3 is drawn at income levels of Rs 100 crores, Rs 200 crores and Rs 300 crores respectively in Figure 9 (A). These curves together with the perfectly inelastic money supply curve *MQ* give us the *LM* curve. The *LM* curve consists of a series of points, each point representing an *interest-income* level at which the demand for money *(L)* equals the supply of money (M). If the income Level is Y_1 (Rs 100 crores), the demand for money (L_1Y_1) equals the money supply *(QM)* at interest rate OR_1. At the Y_2 (Rs 200 crores.) income level, the L_2Y_2 and the *QM* curves equal at OR_2 interest rate. Similarly at the Y_3 (Rs 300 crores) income level, the L_3Y_3 and *QM* curves equal at OR_3 interest rate. The supply of money, the liquidity preference, the level of income and the rate of interest provide data for the *LM* curve shown in Figure 9 (B). Suppose the level of income

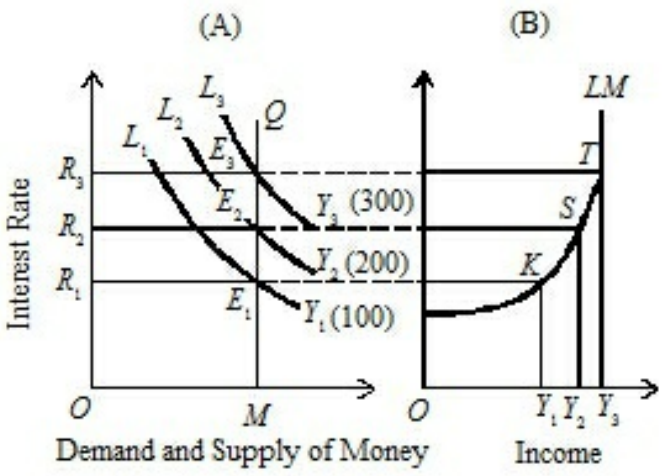


Fig. 9

is Y_1 (Rs 100 crores), as $_s^t$ marked out on the income axis in Figure 9 (B). The income of Rs. 100 crores generates ^{In}a demand for money represented by the liquidity preference curve L_1Y_1 . From the point E_1 where the L_IY_I curve intersects the *MQ* curve, extends a dashed line horizontally to the right so as to meet the line drawn upward from Y_1 at *K* in Figure 9 (B). Points *S* and *T* can also be determined in a similar manner. By connecting these points *K*, *S* and *T* with a line, we get the *LM* curve. This curve relates different income levels to various interest rates, but it does not show what the rate of interest will be.

The *LM* curve slopes *upward* from left to right because given the quantity of money, an increasing preference for liquidity manifests itself in a higher rate of interest. It also becomes gradually *perfectly inelastic* shown as the vertical portion from *T* above on the *LM* curve in Panel (B) of Figure 9. This is because at higher income levels the demand for transaction and precautionary motives increases so that little is left to satisfy the demand for speculative motive out of a given supply of money. We may also note that at the extreme left the *LM* curve is *perfectly elastic* in relation to the rate of interest. This is shown as the horizontal portion of the *LM* curve which starts from the vertical axis in Panel (B) of Figure 9. With the decline in the level of income, the demand for transactions and precautionary motives also declines. Thus a larger amount is available in the form of idle balances but it does not lead to the lowering of the interest rate because we have reached the limit to which the rate of interest will fall. This lower limit to which the rate of interest will fall is the Keynesian *liquidity trap* already explained above in Keynes's theory of interest.

Determination of the Rate of Interest

The *IS* and *LM* curves relate to income levels and interest rates. Taken by themselves they cannot tell us either about the level of income or the rate of interest. It is only their intersection that determines the rate of interest. This is illustrated in Figure 10 where the *LM* and *IS* curves intersect at point E and *OR* rate of interest is determined corresponding to the income level *OY*. The income level and the interest rate lead to simultaneous equilibrium in the *real* (saving-investment) market and the *money* (demand and supply of money) market. This general equilibrium position

persists at a point of time. If there is any deviation from this equilibrium position, certain forces will act and react in such a. manner that the equilibrium will be restored. At the income level OY_t the rate of interest in the real market is Y_1B and it is Y_1A in the money market. When the former rate is higher than the latter rate ($Y_1B > Y_1A$), the businessmen will borrow at a lower rate from the money market and invest the borrowed funds at a higher rate in the capital market. This will tend to raise the level of income to *OY* via the investment multiplier and the equilibrium level of *OR* interest rate will be reached. On the other hand, at the income level OY_2 the rate of interest in the real market is less than the interest rate in the money market *($Y_2C < Y_2D$).* In this situation, the businessmen will try to discharge debts in the money market rather than invest in the capital market. As a result, investment will fall and reduce income by the multiplier to *OY* and the equilibrium rate of interest *OR* will be established,

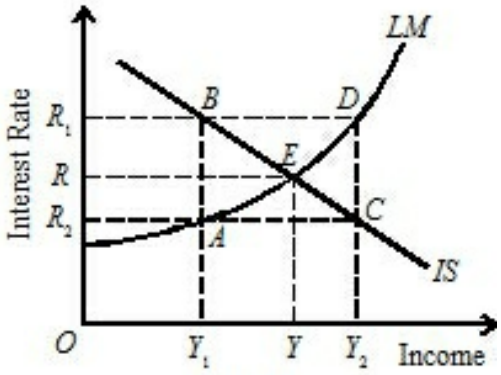


Fig. 10

Shifts or changes in the *IS* curve or the *LM* curve or in both change the equilibrium position and the rate of interest is determined accordingly. These are illustrated in Figure 11. Let *IS* and *LM* be the original curves. They intersect at *E* where *OR* interest rate is determined at *OY* income level. If the investment demand schedule shifts upward, or the saving schedule shifts downward, the curve *IS* would shift to the right as IS_1 curve. Given the *LM* curve, equilibrium will take place at E_1. The rate of interest would be OR_1 and the income level OY_1. If the quantity of money is increased or the liquidity preference cruve is lowered, the *LM* curve would shift to the rights as LM_1. It intersects IS_1

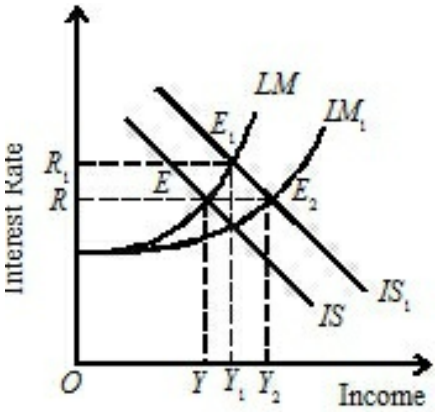


Fig. 11

curve at point E_2. The new equilibrium rate of interest is *OR* and the income level is OY_2. *Thus with a given LM curve, when the IS curve shifts to the right income increases and along with it the rate of interest also rises. Given the IS curve, when the LM curve shifts to the right, income increases but the rate of interest falls. The Hicks-Hansen analysis is thus an integrated and determinate theory of interest in which the two determinates, the IS and LM curves, based on productivity, thrift, liquidity preference and the supply of money, all play their parts in the determination of the rate of interest.*

Its Criticisms

Despite its merits, the Hicks.-Hansen theory of interest rate is not free from certain weaknesses.

1. Static Theory. It is a static theory that explains the short-run behaviour of the economy. Thus it fails to explain how the economy behaves in the long run.

2. Interest Rate not Flexible. The theory is based on the assumption that the interest rate is flexible and varies with changes in *LM* or/and *IS* curves. But it may not always happen if the interest rate happens to be rigid because the adjustment mechanism will not take place.

3. Investment not Interest Elastic. The theory assumes that investment is interest elastic. But if investment is interest inelastic, as is generally the case in practice, then the Hicks-Hansen theory does not hold good.

4. Highly Artificial. According to Don Patinkin, the Hicks-Hansen theory is highly artificial and oversimplified because it divides the economy into real and monetary sectors. In reality, the real and monetary sectors of the economy are so interrelated and interdependent that they act and react on each other.

5. Closed Model. According to Prof. Rowan, the Hicks-Hansen theory is a closed model which does not take into consideration the effect of international trade. This restricts its usefulness for the study of policy problems.

6. Price Level Exogenous Variable. The price level is treated as an exogenous variable in this model. This is unrealistic because price changes play an important role in the determination of income and interest rates in an economy.

Despite these weaknesses, this theory does not undermine the utility *of* the *IS-LM* technique in explaining the determination of interest rate in an economy.

NATURAL AND MARKET RATE OF INTEREST

The Wicksell Theory

Knut Wicksell[5] was the first economist to discuss in detail the relation between natural interest rate and market interest rate. In his book *Interest and Prices,* he uses such phrases as 'ordinary rate', 'the normal rate', and 'the real rate' as synonyms for the natural rate. He defined it in these words: "The rate of interest at which the *demand for loan capital and the supply of savings* exactly agree, and which more or less corresponds to the expected yield on the newly created capital, will then be the normal or natural real rate". It is the rate consistent with a stable money supply and stable prices. On the other hand, the market rate of interest is the money rate prevailing in the loan market. It is the rate of interest charged by banks or lenders. It depends upon the demand and supply of money.

5. *Knut Wicksell, Interest and Prices: A Study of the Causes Regulating the Value ofMoney,* 1898; and *Lectures of Political Economy,* Vol. II, 1906 both translated from Swedish, the former by R.F. Kahn, *1936,* and the latter by *E.* Classen, 1935.

Assumptions

The Wicksell theory is based on. the following assumptions:

1. There is full employment in the economy.
2. Investment is a decreasing function of the rate of interest.
3. Saving is an increasing function of the rate of interest.

Explanation

According to Wicksell, the natural rate is essentially variable. It is partly determined by the demand for loans which, in turn, depends on the expected profitability of new investment. All factors which affect the expected profitability of investment bring changes in the natural rate of interest. They are changes in productive efficiency or technical progress, changes in domestic and foreign demand, changes in the supply of capital, labour and land, etc. But in order to ensure a constant supply of savings, which is the second determinant of the natural rate of interest, the natural rate would have to be the same in all enterprises and in all uses, with the ratio for future land to present land equal to the ratio of future labour to present labour. Any shift in these relationships would change the value of capital and alter the connection between capital goods industries and consumer goods industries.

Wicksell points out that the natural rate is not the same as the market rate. There are disparities between the two rates during the short run which produce changes in the price level. The market rate of interest tends to be sticky and responds slowly to changes in the demand for loanable funds. In the long run, disparities between the two rates automatically generate forces which bring their equality.

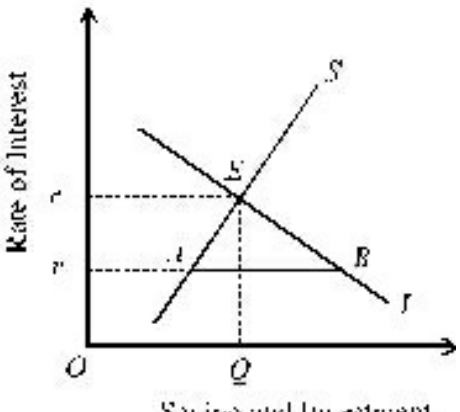


Fig. 12

According to Wicksell, the natural rate diverges from the market rate when the economy is in disequilibrium. As a result, a cumulative process is created whereby discrepancy emerges between the cost of borrowing capital and expected profitability of new investment. A cumulative process is a disequilibrium situation in which net investment is positive and is constantly increasing from period to period. This happens when the market rate is lower than the natural rate. This is shown in Figure 12 where *I* is the investment demand curve or the demand curve for loans and *S* is the supply curve of savings or loanable funds. Suppose *r* is the natural rate of interest and r_1

the market rate of interest. Thus at the market rate of interest r_1, the investment demand (or demand for loans) exceeds the supply of savings by *AB.* It means that bank loans expand and funds are used to increase the demand for investment goods. As a result, the total demand for money exceeds the available supply of money. In a situation of full employment, this raises the demand for goods and services thereby increasing prices. This increased demand for money, in turn, raises the market rate of interest. With the rise in the market rate of interest, money incomes expand and the transactions demand for money increases which reduces the available supply of money for lending purposes. Assuming no further increase in money supply, the money rate of interest comes into equality with the natural rate of interest at point *E* in Figure 12. On the contrary, if the money rate of interest is higher than the natural rate, the demand for bank loans falls, bringing down the market rate of interest till it equals the natural rate.

According to Wicksell, the cumulative process of disequilibrium may also be caused by increase in the demand for loans due to innovations and technical progress which raise the expected profitability of new investment. This is illustrated in Figure 13. The initial monetary equilibrium is shown by the equality of *S* and *I* curves at point *E* where both the natural and market rates are equal at *r* interest rate. The increased demand for loans is shown by the shifting of the investment demand curve upward from *I* to *I,* . This causes the natural rate to rise to r_1 when the *S* and I_1curves intersect at E_1. If the monetary authority does not raise the market rate *r* to the level of the natural rate r_1, the banks increase their lendings at the market rate of interest (r). The increased monetary demand will raise prices. As prices rise aggregate demand would increase along the horizontal axis at the market rate *r.* In terms of Figure 13, at the market rate *r,* the supply of loans by banks is *OQ(= rE)* and with the increase in the demand for investment funds to I_1, banks increase the money supply by QQ_1 (=EA). Thus the horizontal line *EA* is a perfectly elastic supply curve which

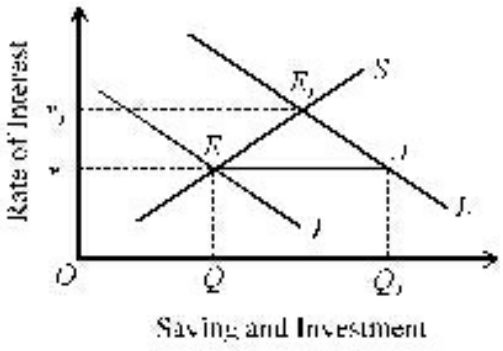


Fig. 13

equals I_1 curve at *A* and the increased investment demand is met by QQ_1 loans from the banks. With the increase in money supply, the demand for capital goods rises which, in turn, raises the demand for goods and services, and consequently raises their prices. This process of money expansion and inflationary rise in prices will ultimately raise the market rate *r* to the level of natural rate of interest r_1. In Wicksell's analysis, business fluctuations are traced to the divergence between natural and market rates of interest. Expansion in the economy takes place when the natural rate is above the market rate, and *vice-versa*. In the cumulative process of upward or downward change, expansions and contractions of bank credit play a crucial role. The cumulative process comes to an end when the two rates are brought into equality. But prices would not be restored to the original level, yet there would be a new equilibrium for the economy where the market rate equals the natural rate.

A Critical Appraisal

Wicksell integrated interest theory, quantity theory, aggregate demand and aggregate supply, and the functioning of a modern banking system. He thus anticipated some of the current notions in monetary theory.

Wicksell stated three conditions in his theory: *First,* the equality between the natural (equilibrium) rate and the market (bank) rate. *Second,* the equality of *ex ante* saving and investment. *Third*, a constant price level. These identical conditions have come to be known as "monetary equilibrium" in modern monetary analysis. It was Wicksell who attributed changes in the price level to discrepancies between aggregate demand and aggregate supply on the basis of relationship between saving and investment. In this way, he formulated an income approach to the problem of money and prices which "contained the embryo of a theory of output as a whole", according to Ohlin.

It was again Wicksell who stressed the importance of the rate of interest in monetary theory. "By concentrating upon the rate of interest", writes Prof. Hansen, "He swept away the narrow foundations of the quantity theory."[6]

Further, Wicksell integrated monetary and non-monetary theories of

interest by emphasising the equilibrium of natural interest rate and market interest rate. He thus paved the way for a determinate theory of interest, as developed by Hicks and Hansen, which is regarded as the modern theory of interest.

Again, Wicksell's cumulative process has come to be known as the *Wicksell Effect.* It emphasises the importance of bank credit-creation on the rate of interest. Mrs Robinson regards the Wicksell Effect as "the key to the whole theory of capital accumulation." Moreover, Wicksell was a quantity theorist who wanted "to arrive at a theory which should be both self-consistent and in full agreement with facts." His analysis of the link between the banking system and short-run inflations during the cumulative process is to be found in the hyper-inflation case studies in Friedman's *Studies in the Quantity Theory of Money.* Similarly the suppressed inflation that followed the post-War years 1945-51 due to the policies adopted by monetary authorities in the USA may also be attributed to the Wicksellian cumulative process. Thus Wicksell was the forerunner of the modern monetary theorists.

Its Defects

But he is not free from critics who have pointed out certain defects in his theory. 1. According to Professor Ackley, "Wicksell's analysis differed from the simple quantity theory only in the *process* by which its results were achieved, not in its results. *In equilibrium*, prices were proportional to the money supply, and both were constant in time. *In equilibrium*, there were no idle balances flowing into the capital market nor additions needing to be made to cash balances in order to finance a larger volume of money transactions."[7]. Thus Wicksell thought of only transactions and precautionary matives neglected the speculative motive to hold money.

---

6. A.H Hansen, *Monetary Theory and Fiscal Policy,p. 71.*

7. G. Ackley *op. cit* p. 161.

2. "Moreover, his work falls short of an adequate theory of income determination. It related to one determinant of income, the investment function," according to Hansen.[8]

3. Hansen further points out that "the multiplier analysis, based on the consumption function was missing. There is, moreover, implicit in much of Wicksell's work an excessively optimistic view with respect to the interest-elasticity of investment. And he saw very dimly the relation of the demand for cash holdings to the rate of interest. He failed to see that in certain situations the investment function may be interest-inelastic while the liquidity preference function may be highly interest-elastic. Wicksell did not clearly understand the conditions under which interest-rate policy becomes futile. Accordingly, he enormously exaggerated the power of the banking system to control, by interest rate manipulations, the flow of aggregate demand and the level of prices:"[9]

Fisher's Analysis

In modern analysis, the *real or natural* rate *of interest refers to the percentage rate paid on borrowed money after making on adjustment for changes in the price level.* On the other hand, *the market or money rate of interest is the percentage* rate paid on *borrowed money.* If a person borrows Rs 10,000 from a bank at 6 per cent, it is the money rate of interest. He will have to return Rs10,600 (Rs 10,000+Rs 600 as interest charges) to the bank after a year. If, however, during this period the price index rises by 1.02, the real rate of interest is reduced. In order to find out the real rate of interest, the sum repaid to the bank is deflated by the price index, i.e. Rs 10,600/1.02=Rs 10,392. It implies that the bank has earned 3.92 per cent as real interest as against the money rate of 6 per cent per annum. If the price index falls, a given money rate of interest will imply a higher real rate of interest.

Following Marshall, Fisher enunciated the proposition that the *money rate of interest is equal to the real rate of interest plus the rate of change in the price level.* Suppose prices are falling at 5 per cent per annum, a zero money rate of interest in this case implies a real rate of 5 per cent. To quote Marshall, a money rate of 5 per cent per annum corresponds to a *real* rate of 15.5 per cent when prices are falling at 10 per cent per annum, i.e., the purchasing power of $ 105 at the beginning of the year is equivalent to that of $ 115.50 at the end of the year. Similarly, a money rate of 5 per cent corresponds to a *negative* real rate of 5.5 per cent when

the annual rate at which prices are rising is 10 per cent.

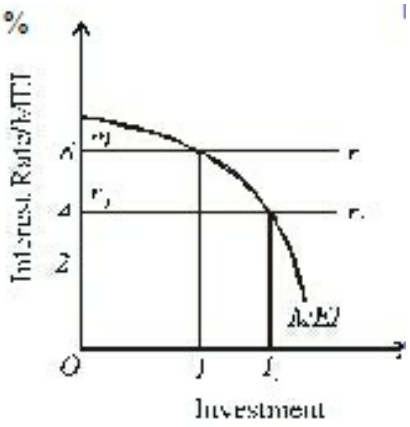


Fig. 14

In case the price level is *constant*, the money rate equals the real rate of interest. To take our previous example, let the money rate of interest and the real rate of interest be 6 per cent with a constant price level. Suppose with a rise in the price level by 2 per cent, the real rate of interest falls to 4 per cent equivalent to the rise in the price level. There is, however, no change in the money rate due to "the money illusion or institutional rigidity in the money market." Given the marginal efficiency of investment, a fall in the real rate of interest will encourage investment and *vice-versa.* This is illustrated in Fig 14. Money and real interest rate are taken on the vertical axis along with the marginal efficiency of investment. The amount of investment is shown on the horizontal axis. *MEI* is the curve of marginal efficiency of investment. With a constant price level, *mr* represents money as well as real rate, of interest which is 6 per cent per annum. At this rate, *OI* investment is taking place. With the rise in the price level by 2 per cent, there is an equivalent fall in the real rate, as depicted by the new interest line r_1 r_2 at *4* per cent interest rate. Now the *mr* line represents only the money rate of 6 per cent. A fall in the real rate of interest has tended to encourage investment from *OI* to OI_1. Converse is true when the price level falls.

---

8. A.H. Hansen op. *cit., p,* 89.

9. *Ibid.,* p, 93.

EXERCISES

1. Discuss the role of the real forces of thrift and productivity in determination of the rate of interest.
2. In what respects is the loanable funds theory of interest an improvement on the classical theory?
3. Interest theory has been confused by a failure to distinguish clearly

between the money rate and the natural rate of interest." Discuss.

4. Critically discuss the Keynesian theory of liquidity preference as an explanation of the determination of the interest rate.
5. Compare the classical and the Keynesian theories of the rate of interest.
6. How are rate of interest and national income determined simultaneously in an economy?
7. Explain the Hicks-Hansen reconciliation of the stock and flow theories of interest rate.
8. Discuss Wicksell's theory of interest rates.
9. Write a note on Fisher's theory of real and market rate of interest.

CHAPTER 34

TERM STRUCTURE OF INTEREST RATES

MEANING

The term structure of interest rates refers to the relationship between market rates of interest on short-term and long-term securities. It is the interest rate difference on fixed income securities due to differences in time of maturity. It is, therefore, also known as time-structure or maturity-structure of interest rates which explains the relationship between yields and maturities of the same type of security.

If two securities are identical in every respect except maturity, it is likely that they will sell in the market at different prices (or yields or interest rates). Generally, their prices will change in the same direction. If the short-term securities rise in price, the long-term securities will also rise in price. People generally hold both short-term securities and they adjust their holdings of securities depending on the relative yields. Usually the long term securities tend to fluctuate more in price than the short-term securities, even though their yields do not fluctuate as much.

The relationship between yields and terms to maturity is depicted graphically by a yield curve. Figure 1 shows three yield curves. When short-term interest rates are above long-term interest rates, the yield curve slopes downward, as the curve *FF*. When short-term interest rates are below long-term interest rates, the yield curve slopes upward, as the curve *RR*. When the short-term yields equal long-term yields, the yield curve is flat, as the curve *HL*.

*Students should draw Fig. 1 here.

FACTORS DETERMINING THE TERM STRUCTURE OF INTEREST RATES

What determines the shape of the yield curve? There are three factors which determine the term structure of interest rates. They are risk preference, supply and demand of securities, and expectations and uncertainty. These factors determine whether short-term interest rates are above or below long-term interest rates. We discuss these factors as under:

1. Risk Preference. Long-term security prices are sensitive to changes in interest rates because the chances to default are higher on long-term securities as compared to short-term securities. Therefore, lenders prefer to lend for short-term, if short-term and long-term securities have identical yields. This would push up short-term prices of securities and bring down their yields. Hence the yield curve slopes upward. On the other hand, borrowers prefer to borrow for long period because they will not have to worry about rising interest rates or to renew their loans frequently. They are, therefore, willing to pay more for long-term securities as compared to short-term securities. This will also cause the yield curve to slope upward. Thus the preferences of both lenders and borrowers lead to low rates on short-term securities and high rates on long-term securities, thereby bringing about an upward sloping yield curve.

2. Supply-Demand Conditions. When the supply of short-term securities falls and that of long-term securities rises, the short-term interest rate comes down and the long-term interest rate is pushed up. The yield curve is upward sloping and vice versa. If the demand for securities is more in the short-run market and the supply is more in the long-run market, this will lead to high short-term and low long-term interest rates, and the yield curve will be downward sloping. The opposite supply-demand conditions will lead to an upward sloping yield curve.

3. Expectations and Uncertainty. Other factors affecting the yield curve are expectations and uncertainty. The expectation of the rise in the long-term interest rate explains that the short-term interest rate remains much below the long-term interest rate for any length of time. This produces an

upward sloping yield curve. Further, certain risks and uncertainties may lead to the same results. For instance, if people expect war, social disturbances, political upheavals, uncertainties, inflationary pressures, etc., they will not purchase long-term securities except at a low price or low current yield.

THEORIES OF TERM STRUCTURE OF INTEREST RATES

Many theories have been put forth by economists to explain differences in the structure of interest rates on short-term and long-term securities. They are discussed as under :

The Expectations Theory

The expectations theory regards future interest rates as the principal determinant of the present structure of interest rates. The theory originated with Irving Fisher, was perfected by Hicks in his *Value and Capital*, and is closely identified with Lutz.[1]

Its Assumptions

The expectations theory is based on the following assumptions:

1. All investors have definite expectations with respect to future short-term interest rates, and these expectations are held with complete confidence.

2. The objective of investors is to maximise expected profits, and they are prepared to transfer funds freely from one maturity to another in order to achieve this objective.

3. There are no costs associated with investment and disinvestment in securities. It means that there are no transaction costs.

4. The short-term and long-term interest rates are adjusted for any differences due to risk and liquidity.

5. ‘Safe’ securities of various maturities are perfect substitutes in the

portfolios of investors.

6. Investors are profit maximisers who hold such financial assets in their portfolios which maixmise return over a period they are held.

7. All investors hold with certainty the same expectations of how future rates are going to behave.

Explanation

Given these assumptions, the theory states that the long-term interest rate at any point in time represents an average of expected short-time interest rates. Suppose a long-term security maturing after three years sells at the short-term interest rate of 2 per cent in the first year, and that the expected short-term interest rates in the second and third years are 3 per cent and 4 per cent respectively. The long-term interest rate on this security will be the average of the short-term interest rates over the three years, that is 3 per cent (=2 per cent + 3 per cent+4 per cent=9 per cent/3). If the interest rate on the short-term security for the first year is expected to decline by 1 per cent the long-term yield on the three-year security will be 2.67 per cent (=1 per cent+3 per cent+4 per cent=8 per cent/3). On the other hand, if the interest rate is expected to increase by 1 per cent on the three-year security, then the long-term yield will be 3.33 per cent (=2 per cent+3 per cent+5 per cent=10 per cent/3).

The expectations theory holds that differences in yields on securities of different maturities are due to the fact that the market expects the interest rates on different securities to be the same over an equal period of time. If this is not the case, the investor will buy security of one maturity by selling security of another maturity that he expects to provide him the highest yield. It is this “arbitrage procedure” (buying security of one maturity and selling of another maturity) which brings the equality of expected yields of different maturities.

According to Warren L. Smith, “Investors generally have *repressive interest rate expectations*. That is to say, at any particular time they have an opinion regarding the level of interest rates they regard as normal, and as short-term rates rise above or fall below this level, they expect them to

regress back towards this normal level. Thus, as rates rise above normal, investors expect them to fall; and as rates fall below normal, investors expect them to rise." This relationship implies that "(*a*) when short-term rates are expected to fall, current short-term rates will be above long-term rates and the yield curve will be negatively sloped, and (*b*) when rates are expected to rise, current short-term rates will be below long-term rates and the yield curve will be positively sloped. By the same reasoning, when the short-term rate is at approximately the level judged to be normal and is expected neither to rise nor to fall, rates for all maturities will appoximate a horizontal line."[2]

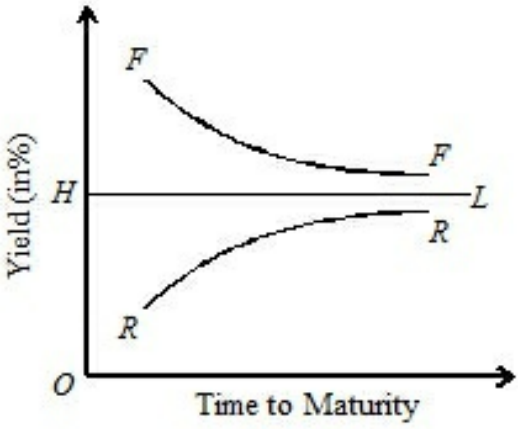


Fig. 1.

---

1. F.A. Lutz, "The Sturcture of Interest Rates", *Q.J.E.*, November 1940, reprinted in W.Fellner and B.F. Hally (Eds.), *Readings in the Theory of Income Distribution*, 1946, Ch. 26.

These relationships between short-term and long-term interest rates are illustrated in Figure 1. where time to maturity, short or long-term, is taken on the horizontal axis, and yield or interest rate is taken on the vertical axis. In the first case, when the current short-term rates are above the long-term rates and are expected to fall, the yield curve *FF* is negatively sloped. In the second case, when the current short-term rates are below the long-term rates, and are expected to rise, the yield curve *RR* is positively sloped. In the third case, when the current short-term rates equal the long-term rates and the short-term rates are expected neither to rise nor to fall, the yield curve *HL* is a horizontal line.

Prof. Lutz calls these yield curves as *expectations curves* which represent "the line up of the'subjective' long rates which correspond to people's expectations after the latter have changed."

The key to the expectations theory is that both short and long-term securities are perfect substitutes for each other in the portfolios of investors. If the supply of long-term securities goes up and that of short-

term securities goes down, it makes no difference as far as yields are concerned. Long-term securities will be exchanged for short-term securities with no change in yields as long as expected future short-term rates are unchanged.

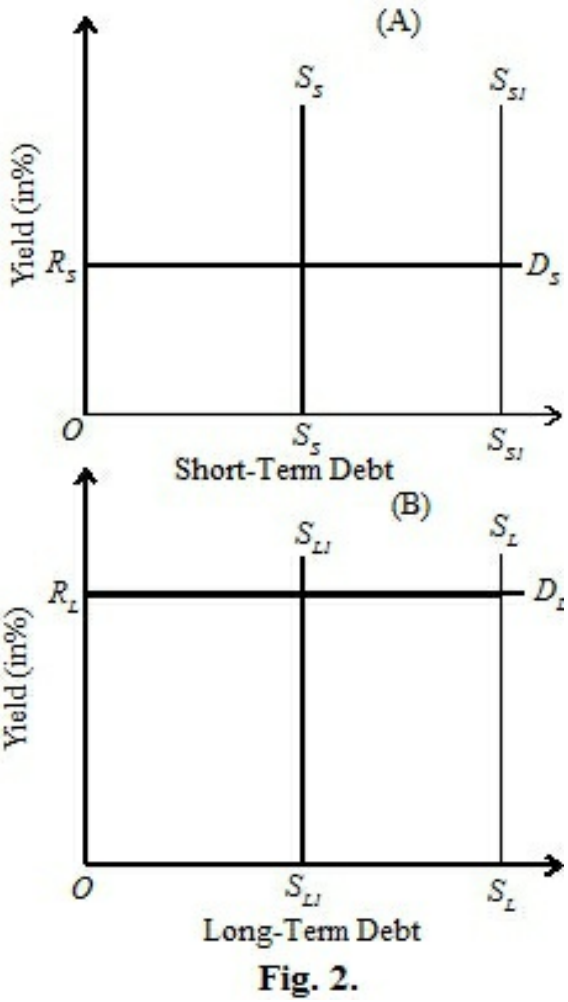


Fig. 2.

Its Policy Implications. The policy implications of this theory are that if the government wishes to replace a given amount of long-term debt with an equivalent amount of short-term debt, such a policy will have no impact on the structure of interest rates. This is because both long-term and short-term debts are regarded as perfect substitutes in the portfolio of investors. This is shown in Figure 2 (A) and (B). Panel (A) depicts the demand and supply of short-term debt and Panel (B) that of long-term debt. Each type of debt is measured on the horizontal axis and yield on the vertical axis. The demand curves D_S and D_L are drawn with infinite elasticity because both debts are perfect substitutes. The supply cuves S_S and S_L are drawn perfectly inelastic becuase the quantities of short-run and long-term debts are fixed. Assume that the long-term interest rate (yield) lies above the short-term interest rate, ie., $OR_L > OR_S$. This is because the future short-run interest rate is expected to be higher than the current short-run rate. If the government wishes to redeem (retire) $S_L – S_{L1}$ of long-term debt and replaces it with $S_S – S_{S1}$ of short-term debt, this is possible without any change in the relative yields of the two kinds of debt. These remain the same at OR_L and OR_S repectively.

Again, as all types of short-term and long-term bonds (securities) are perfect substitutes, therefore the central bank cannot influence the term

structure of interest rates as long as it does not affect the expectations of borrowers and lenders.

---

2. Warren L. Smith, *Macroeconomics*, 1970, p. 283.

Its Criticisms

However, the expectations theory has been criticised on several points:

1. Lenders may have expectations about long-term interest rates that may be independent of their expectations about short-term interest rates.

2. The theory presupposes that investors can make long-term expectations about short-term interest rates. But it is doubtful if such predictions can be made accurately.

3. Critics doubt the efficacy of changes in the central bank discount rates to influence the long-term interest rates. For instance, a reduction in the discount rate can bring a fall in the long-term rates only if the expectation is generated that short-term interest rate will remain low. This will prevent the discount rate being changed very often by the central bank.

4. If open market operations which influence the slope of the yield curve are not successful, the expectations theory fails. Suppose the central bank tries to keep long-term rates higher than short-term rates by supplying short-term market with large funds. In this situation, investors expecting that they will gain over the long period would shift from short-term to long-term securities. This would tend to equalise the short-term and long-term rates and the yield curve would be horizontal like *HL*, rather than like *RR* in Figure 1.

5. The assumption of the theory that investors hold with certainty expectations of future short-run interest rates is not correct. This is because expectations of people for short-run interest rates are occasionally certain.

6. This assumption is also away from reality that the expectations of borrowers and lenders are similar. In fact, the expectations of borrowers

and lenders regarding short-term and long-term rates are quite different from each other.

7. The theory fails to explain how expectations relating to future short-term interest rates are formed.

8. It is also wrong to assume that transaction costs are zero. In reality, borrowers and lenders are required to incur transaction costs every time when they buy and sell securities.

The Segmented Markets Theory

The segmented markets theory is known by various names such as institutional, hedging or segmentation. According to this theory, investors are much averse to risk. So they hedge against risks by matching the maturity of their assets with that of their liabilities. If the maturity of an investor's assets is longer than that of his liabilities, he incurs a capital loss when he is forced to sell his assets before they are due for redemption. On the other hand, if the maturity of an investor's assets is shorter than that of his liabilities, he runs the risk of income loss. In order to avoid the two kinds of risks, investors match the maturities of their assets and liabilities.

Its Assumptions

This theory is based on the following assumptions:

1. Assets of different maturities are imperfect substitutes with each other.

2. Markets for assets of different maturities are divided into separate markets.

3. Interest rates for one type of asset in every market are determined by their demand and supply which, in turn, affect the yield to maturity.

4. There is uncertainty about the behaviour of interest rates in future.

The Theory

The segmented markets theory holds that short-term and long-term interest rates are determined in several separated or segmented markets. Some investors prefer short-term securities, while other investors, such as insurance companies, prefer long-term securities. Thus securities of different maturities are imperfect substitues for buyers and sellers of securities in the market.

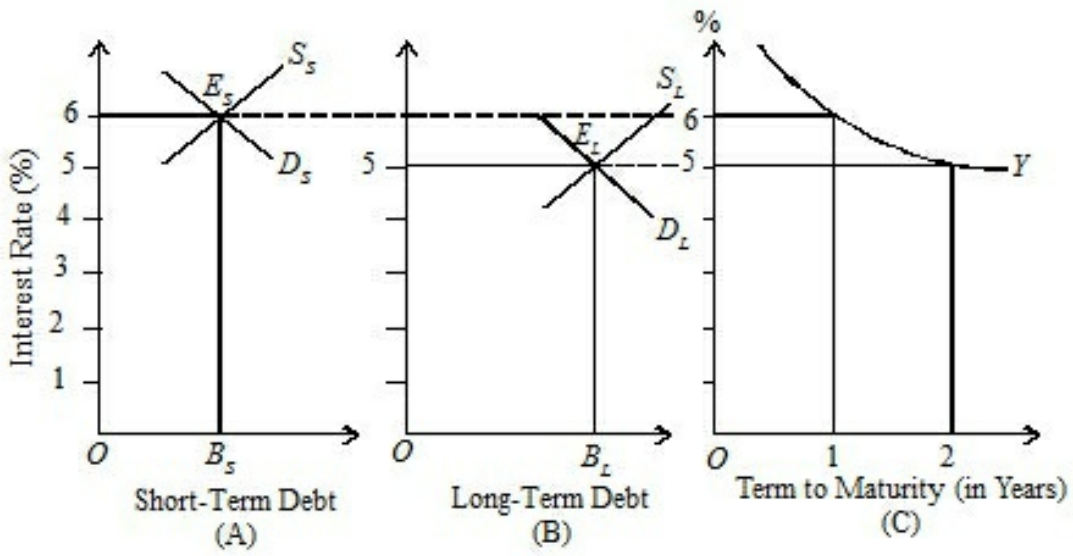


Fig. 3.

Consequently, the yield curve is the result of several demand and supply curves for securities of different maturities. Given the demand for securities, if the supply of short-term securities is less than the demand for long-term securities, the short-term interest rate will be higher than the long-term interest rate. In this situation, the yield curve will slope downward to the right, as shown by the curve Y in Fig. 3 (C). In Panel (A) of the figure, $D_S$and $S_S$are the demand and supply curves of short-term debts respectively which are in equilibrium at point E_S. Thus they determine 6% equilibrium interest rate on short-term securities. In Panel (B), $D_L$and $D_S$are the demand and supply curves respectively of long-term debts which determine 5% equilibrium interest rate at point $E_L$on long-term securities. These interest rates provide the downward sloping yield curve *Y* in Panel (C).

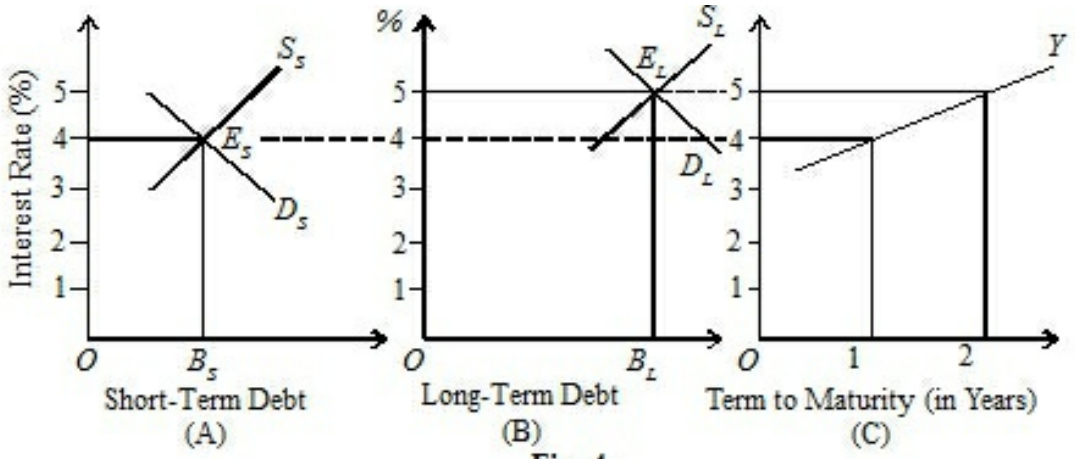


Fig. 4.

On the contrary, given the demand for securities, when the supply of short-term securities is greater than the demand for them, the short-term interest rates will be lower than the long-term interest rates. In such a situation, the yield curve will slope upward to the right, as shown in Fig. 4 (C). In Panel (A) of the figure , $D_S$and $S_S$curves determine 4% short-term equilibrium interest rate at point $E_S$on short-term assets. Similarly in Panel (B), $D_L$and $S_L$curves determine 5% long-term equilibrium interest rate at point E_L. These interest rates provide the upward sloping yield curve Y in Panel (C). Thus, according to the segmented market theory, the market rates on securities of different maturities are determined by separate conditions of demand and supply in each maturity. Across the maturity spectrum, there are a number of separate markets.

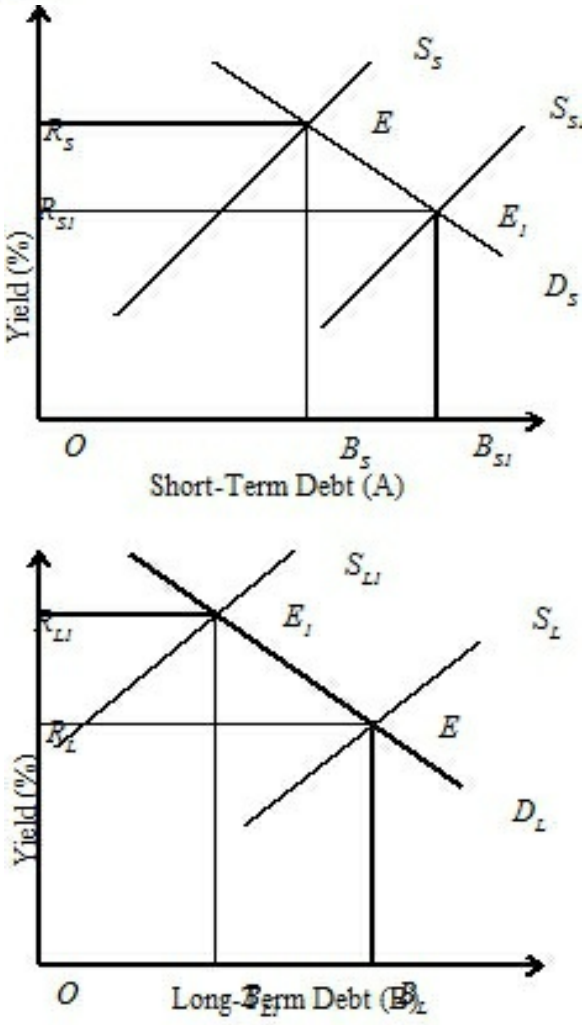


Fig. 5.

Its Policy Implications

The policy implications of this theory are that if the government wishes to replace a given amount of long-term debt by a short-term debt, it will be successful in twisting the structure of interest rates. In this theory, $D_S$and $D_L$are the demand curves for long-term and short-term debts respectively which are less than perfectly elastic with respect to yield rates, as shown in Figure 5 (A) and (B). This is because the two types of debt are not perfectly substitutable. Suppose that the government wants to retire B_L–$B_{L1}$of long-term debt and replace it with B_{S1}–$B_S$short-term debt. Notice that the spread between the rates has been reduced. When the government substitutes long-term debt by short-term debt, the supply of long-term debt is reduced from $S_L$to $S_{L1}$and the long-term interest rate rises from $R_L$to R_{L1}, as shown in Panel (B) of the figure. On the other hand, the supply of short-term debt increases from $S_S$to S_{S1} which brings a fall in the short-term interest rate from R_S to R_{S1}, as shown in Panel (A) of the figure. Notice that the fall in the short-term interest rate is less than the rise in the long-term interest rate : R_S–$R_{S1} < R_L$–R_{L1}.

Another implication of this theory is that the central bank can affect the yields to maturity of securities or the term structure of interest rates by permitting the relative supplies of long-term and short-term securities.

Again, the central bank cannot affect the long-term interest rate by changing only the supply of short-term securities.

Its Criticisms

The segmented market theory has been criticised on the following grounds :

1. This theory explains that changes in preferences for securities of different maturities will alter in the form of the yield curve. But it does not explain changes in the structure of yields.

2. This theory is based on the assumption that short-term and long-term interest rates are not related to each other. But empirical evidence does not support it. It shows that short-term and long-term interest rates move

together. When short-term interest rates fall or rise, the long-term interest rates move in the same direction.

Conclusion. Despite these criticisms, the segmented markets theory is supported by institutional practices. Accordingly, commercial banks which place emphasis on liquidity, deal in short-term securities, and insurance companies with long-term securities. Similarly, inventories are financed with short-term loans, and purchases of houses with long-term mortgages.

Its Superiority Over Expectations Theory

The segmented market theory is superior to the expectations theory on the following counts :

1. The segmented market theory is superior to the expectations theory because it does not assume the unrealistic assumption of the expectations theory that short-term and long-term securities are perfect substitutes. The risks in short-term securities are less than those in long-term securities. If a person sells his security before its maturity and the interest rate is more than expected, the price of the long-term security will be less as compared with the short-term security. Thus, short-term and long-term securities are not perfect substitutes.

2. The segmented market theory is also superior to the expectations theory because it rejects the assumption of the latter that the future interest rates are known with certainty. In reality, they are uncertain. Due to large price-changes of long-term securities, there is much uncertainty in holding them. On the other hand, there is great uncertainty of future yields in holding short-term securities.

3. Another reason of the superiority of segmented market theory over the expectations theory is that it does not explain the term structure of interest rates on the basis of the average of expected short-term interest rates. Rather, the segmented market theory determines both the short-term and long-term interest rates in the form of demand and supply of a particular security, as happens in reality in a financial market.

4. As against the expectations theory, the segmented market theory does not explain a unique relation between short-term and long-term interest rates. In reality, the behaviour of short-term and long-term interest rates depends on the relation between money market and bond market. Short-term and long-term interest rates in both markets are determined by the demand and supply of each type of security.

5. The segmented market theory is also superior to the expectations theory because it is supported by institutional practices. Accordingly, commercial banks which place emphasis on liquidity deal in short-term securities, and insurance companies with long-term securities. Similarly, inventories are financed with short-term debts and purchases of houses with long-term mortgages.

The Substitutability Theory

The substitutability theory holds that short-term and long-term securities are substitutes for borrowers and lenders. When buyers and sellers of securities are engaged in arbitrage and switching operations, they tend to eliminate discrepancies between long-term and short-term interest rates in the short run. For such operations, the theory assumes optimising behaviour on the part of buyers and sellers, and relatively free and unrestricted markets.

Given these assumptions, prices of short-term securities move in the same direction as the long-term securities. For instance, a fall in the price of short-term securities (or a rise in the short-term interest rates) will be followed by a fall in the prices of long-term securities (or a rise in the long-term interest rates), and vice versa.

The forces which lead to such parallel moments between short-term and long-term interest rates are (*a*) the substitutability of alternative investment opportunities on the part of lenders and the alternative means of financing on the part of borrowers, and (*b*) by tendency for changes in credit and monetary conditions to have a simultaneous impact on the financial market.

The substitutability theory is important for conducting monetary policy

when by arbitrage and switching operations, changes are transmitted from one sector of the financial market to the other by the central bank.

However, it is doubtful if investors will readily switch over from short-term to long-term securities. Therefore, the perfect substitutability assumed between the two types of securities breaks down.

The Keynesian Theory

There are two theories associated with Keynes. The *first* is known as the *fluidity theory* which Keynes put forth in his *Treatise on Money*, and the *second* is called the *psychological theory* which he propounded in his *General Theory*. The fluidity theory holds that there is correspondence, both in direction and timing, of movement in interest rates with short-term interest rates moving proportiontely more than long-term interest rates. Further, any action on the part of the central bank to influence short-term interest rates is readily transmitted to long-term interest rates. But Keynes himself admitted that during a slump in security prices, the central bank's action to influence short-term securities through open market operation may not help in reducing yield on long-term securities.

The psychological theory holds that the long-term interest rate is a highly psychological phenomenon dominated by short-run expectations about its future level. The long-term interest rates are sticky in the downward direction even when short-term interest rates decline substantially. This is due to the psychology of investors. If they believe that the long-term interest rates have reached the "irreducible conventional minimum" and the next change would be in the upward direction, they will not deal in long-term securities for fear of capital loss. This happens during a depression when liquidity preference is perfectly interest elastic. All efforts on the part of the central bank to purchase short-term securities in the open market will fail to persuade investors to buy long-term securities.

These Keynesian theories of term-structure of interest rates have been neglected and attention has focussed on the expectational theory and the segmented markets theory which have been empirically tested.

The Liquidity or Risk Premium Theory

A variant of the Expectations Theory and the Keynesian Fluidity Theory is called the Liquidity or Risk Premium Theory. The theory rejects the view that short-term and long-term securities are comparable except for maturity. But it accepts the view that yields on various maturities are related to each other by the expectations of long-run and short-run rates. Thus it adds to the expectations theory that lenders and borrowers differ in their attitudes to short-term and long-term securities. Lenders prefer to lend for short-term and borrowers prefer to borrow for long-term. This is because the return on short-term securities is certain but that on long-term securities is uncertain on account of uncertainties of future interest rates. The uncertainty is greater, the longer the maturity of securities. Thus risks associated with long-term securities are greater. Under the circumstances, lenders prefer the safer short-term securities. To induce the market to hold the long-term securities supplied by long-term borrowers, the expected return on them must exceed that on short-term securities by a premium which is called an *expected risk or liquidity premium*. Thus long-term securities being more risky would command larger liquidity premiums.

As long-term securities carry higher risk premium, it follows that interest rates increase as one moves from short-term to long-term securities. The higher interest rates on longer-term securities are determined by two components: (a) the expectations component, and (b) the liquid premium component. The *liquidity premium* component will always lead to higher interest rates on long-term securities than on short-term securities. This tendency has been called by Hicks the “constitutional weakness” or the “congenital market weakness” by some other supporters of the theory. This tendency of higher interest rates may be counteracted partially or even fully by the *expectations component* which is assumed as given.

Prof. Newelyn in his *Theory of Money* has modified this theory by dropping the expectations component and introducing the concepts of encashment and requirement periods. The concept of *encashment period* is associated with the *lender*. The lender calculates the period for which he is prepared to lend his funds and at the end of which he will encash them. This is the encashment period. If he lends his funds for a period shorter than the encashment period, he is not certain about his interest income which he may be able to get for the remaining period. A part of the

interest income will depend on the future behaviour of interest rates. Thus the lender runs an “income risk”, accoding to Newelyn, which the lender can avoid if the duration of the loan is equal to the encashment period. On the other hand, if the loan is longer than the encashment period, he runs a capital risk. This may be due to differences in the market price of a loan from its face value. It follows that the lender can avoid both the risks if he lends his funds exactly for the encashment period. So long as the duration of the loan does not exceed the encashment period, the lender is not to be paid any liquidity premium to induce him to part with his funds.

The concept of the *requirement period* is related to the *borrower* who calculates the period for which he requires the loan and at the end of which he will repay it. If the borrower borrows for a period equal to his requirement period, the interest cost of the loan is predetermined and there is no uncertainty about this. When the borrower borrows for the duration of the requirement period, he avoids the income risk which is the interest risk cost. If he borrows for a shorter period than the requirement period and thinks of borrowing for the remaining period later on, he runs the risk of incurring a different interest cost, and thereby incurs an income risk. If he borrows for a longer period than the requirement period, he may pay a price different from the face value of the debt for the excess period. In this case, he runs the capital risk. The borrower can avoid both the risks if he borrows exactly equal to the requirement period.

We find that both the lender and the borrower can avoid the income and capital risks if the maturity of the loan coincides with the encashment and the requirement period repectively. Newelyn’s analysis is based on the presumption that the average requirement period of the borrowers exceeds the average encashment period of the lenders. This lures the lenders to lend for a period longer than the encashment period and thereby run the capital risk. As a result, longer maturity loans would be carrying a premium against the capital risk. The premium would be larger as we move from shorter to longer maturities which would move the interest rates upward. Thus long-term interest rates tend to be higher than short-term interest rates and the market demands a premium called *liquidity premium* to hold long-term securities. According to Newelyn, “This can be done by relating the maturity of existing assets to the maturity which

would be required if all assets were to mature at the date at which the holders anticipate that ancashment be required."

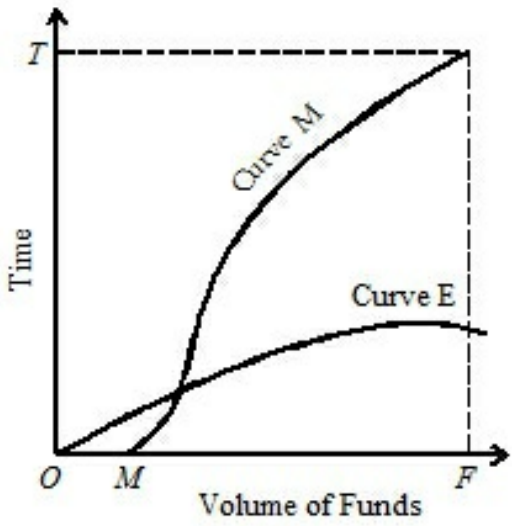


Fig. 6.

This is explained in Figure 6. The encashment curve *E* shows the cumulative total funds measured on the horizontal axis increasing to *F* as the period within which encashment is expected to increase to time *T*. Similarly, the maturity curve *M* shows the cumulative total assets increasing to *F* as the period within which their maturity increases to time *T*. The curve *M* rises from a positive value of *M* at zero life to maturity, indicating the existence of money and quasi-money (*M*). Both the curves represent the same total amount of funds (*F*), reflecting that the total funds must be distributed among the existing stock of assets. But they explain the different distributions of two characteristics of this total of funds: (i) the encashment period of the funds and the length of life to maturity of the assets in which the funds must be held; and (ii) the time period *T* is assumed to be sufficiently long to include all bonds and the maximum encashment period.

The area under the maturity curve *M* will change inversely with changes in the degree of maturity of the existing stock of assets. On the other hand, the area under the encashment curve *E* will change inversely with changes in the degree of maturity consistent with zero capital risk for the investor. Therefore, the general level of interest rates will be a function of the difference between the two areas because this difference measures the extent to which investors must be induced to accept the risk of capital loss if the demand for maturity is to be equated with the given supply of funds.

Conclusion. This theory is more realistic than the expectations theory because it takes into consideration the risk involved by investors of securities of different maturities. Rather, it is an improvement over the other term structure theories and is the most acceptable among the

majority of economists.

The Preferred Habitat Theory

Modigliani and Sutch[5] have propounded the *Preferred Habitat Theory* of the term structure of interest rates. It combines the main features of both the expectations and segmented market theories. According to this theory, investors have a preference for securities of a given term and they want to choose them according to their expected yield. But they will be willing to purchase securities of some other term by substituting them for securities of a preferred term. They will do so if they are compensated by the term premium. The term premium is a compensation or an additional yield which induces investors to purchase securities with a different term to maturity than their preferred term.

According to this theory, different investors have different habitats.* Suppose an investor has *n* period habitat which means that he has funds which he does not need for *n* periods and keeps them invested in securities for *n* periods. He knows exactly the result of investment as measured by terminal value of his wealth. If he invests for a short-period, his outcome is uncertain, as it will depend on the future course of the short-term rates. Moreover, he will have to incur greater transaction costs. Thus, if he has risk aversion, he will prefer to hold long-period securities. On the contrary, if the average short-term rates exceed the long-term rate by an amount sufficient to cover extra transaction costs and to compensate him of the extra risk of holding short-period securities, he will stay short. But risk aversion should not lead him to prefer to stay short.

If the investor invests in maturities longer than *n* period habitat, he would face uncertainty as to the price he can get for his unmatured securities. Thus risk aversion should lead him to hedge against risks by continuing to remain in the same maturity habitat unless other longer or shorter maturities offer an expected premium sufficient to compensate for the risk and cost of moving out of his habitat. Similar considerations apply to the borrowers in the market.

Under this theory, if the *n* period demand for funds exceeds the funds with *n* period habitat, there will arise a premium in the *n* period maturity and

vice versa. Such premiums have a tendency to bring about shifts in funds between different maturity markets in two ways: *first*, through the speculation of investors who are tempted out of their natural habitat by the lure of higher expected returns, and *second*, through arbitrage by intermediaries by borrowing in the maturity range where the expected return is low, and by lending where the expected return is high.

Thus the preferred habitat theory tells that an investor chooses securities on the basis of both the expected yield on securities and his preference for securities with a particular maturity. He will purchase a security of *undesirable* term to maturity only if he receives a term premium.

Suppose an investor prefers a one-year security because he needs funds at the end of the year. If he purchases a two-year security and sells it at the end of one year before its maturity, he incurs a risk. If the interest rate rises during this period, he will be forced to sell the security at a lower price (because security prices and interest rates are inversely related). In addition, he will have to pay commission to the broker in order to sell the security before maturity of two years. Therefore, the investor must recieve a term premium to pruchase a long-term security when his preference is for a short-term security.

---

5. F. Modigliani and R. Sutch, "Innovations in Interest Rate Policy", *A.E.R.*, May 1966.

*In this theory, 'habitat' refers to short-term security (habitat) or long-term security (habitat)

The preferred habitat theory explains the *shapes* of the *yield curve* in a better way than the expectations hypothesis and the segmented market hypothesis. It explains not only upward sloping yield curves but also horizontal and downward sloping yield curves. It also explains why the yield curves are usually upward sloping rather than downward sloping. When the term premium is positive and there are expectations of short-term interest rates to *rise* by a *large amount*, the yield curve will be upward sloping, as shown in Fig. 7 (A). Again, when the

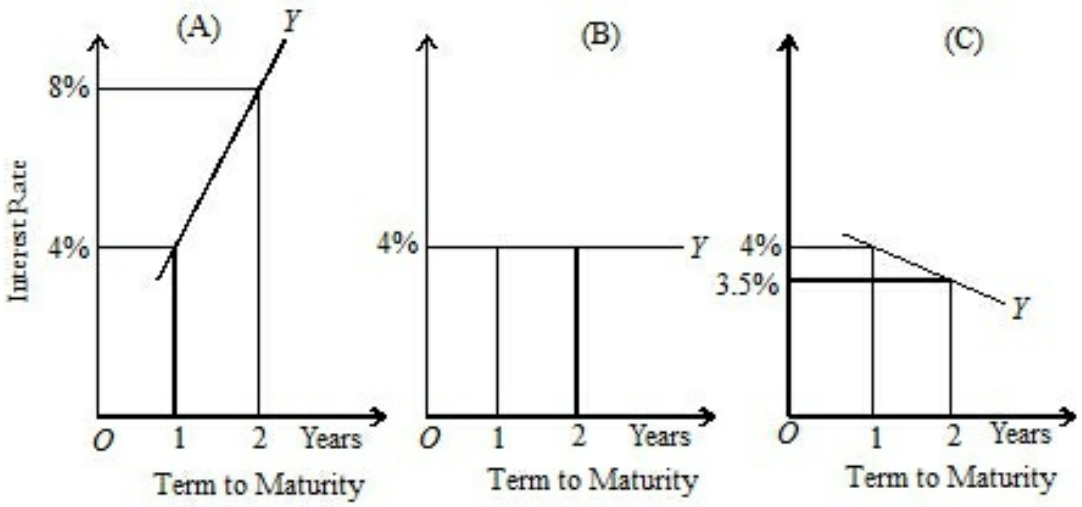


Fig. 7.

term premium is positive and the short-term interest rates are expected to *rise by a small amount*, the yield curve will be relatively horizontal (flat), as shown in Panel (B). If the term premium is positive and the short-term interest rates are expected to *fall* considerably, the yield curve will be downward sloping, as shown in Panel (C). If, however, the short-term interest rates are expected to remain constant or fall a little, the yield curve will slope upward on account of the term-premium on long-term securities.

Conclusion. The preferred habitat theory points out that expecations, risk premium and market segmentation all play a part in determining the term structure of interest rates. If lenders and borrowers in the capital market are not rigidly tied to market segments, but simply have preferred habitats, then expectations play their part in determining interest rates which are not completely independent. Other things being equal, lender's preference for liquidity will have a tendency for long-term rates to be above short-term rates.

In reality, the expectations and segmented market theories are the exptreme versions of the preferred habitat theory. In the expectations theory there is no term premium, that is, it is zero. It is simply the average of expected rates during the term of the long-term security. Investors do not require any premium to hold securities of different maturities. But in the segmented market theory, the term premium is *infinite*. There is no

term interest rate which can be offered to an investor to induce him to purchase a security whose maturity does not match his preferences.

Empirical evidence on the habitat theory suggests that the link between expected short-term rates and long-term rates is quite strong, thus justifying the expectations theory. Further, the expected return from long-term securities tends to exceed the short rate by a positive premium. The existence of such a positive premium would indicate a systematic tendency for the primary supply of funds to exeed the primary demand for them in the short period market, and to fall short of the primary demand in the long period market. Under these conditions, the size of the premium on long-term securities would depend on the facilities for effective arbitrage operations.

EXERCISES

1. What do you understand by the term structure of interest rates? Explain the factors which determine short-term and long-term interest rates.
2. What do you mean by the structure of interest rates? Discuss the relationship between short-term and long-term rates of interest.
3. What do you mean by the term structure of interest rates? How do you explain the relationship between interest rates of different maturities in a given time structure?
4. Critically discuss the expectations and the segmented markets theories of term structure of interest rates.
5. Write notes on: (i) Risk Premium Theory, (ii) Preferred Habitat Theory.
6. Explain the pereferred habitat theory. How is it superior to the expectations and segmented market theories of term structure of interest rates.

CHAPTER 35

THE REAL BALANCE EFFECT AND PIGOU EFFECT

PATINKIN'S INTEGRATION OF MONETARY THEORY AND VALUE THEORY : THE REAL BALANCE EFFECT

Don Patinkin in his monumental work *Money, Interest and Prices* [1]criticises the Cambridge economists for the homogeneity postulate and the dichotomisation of goods and money markets and then reconciles the two markets through the real balance effect.

The *homogeneity postulate* states that the demand and supply of goods are affected only by relative prices. It means that a doubling of money prices will have no effect on the demand and supply of goods. Mathematically, the demand and supply functions for goods are homogeneous of degree zero in prices alone. Thus this homogeneity postulate precludes the price level from affecting the goods market as well as the money market. Patinkin criticises this postulate for its failure tohave any determinate theory of money and prices.

Another closely related assumption which Patinkin criticises is the *dichotomisation of the goods and money markets* in the neo-classical analysis. This dichotomisation means that the relative price level is determined by the demand and supply of goods, and the absolute price level is determined by the demand and supply of money. Like the homogeneity postulate, this assumption also implies that the price level has absolutely no effect on the monetary sector of the economy, and the

level of monetary prices, in turn, has noeffect on the real sector of the economy.

1, 2nd Edition, 1965.

After condemning the neo-classical assumptions outlined above, Patinkin intergrates the money market and the goods market of the economy which depend not only on relative prices but also on real balances. Real balances mean the real purchasing power of the stock of cash holdings of the people.

When the price level changes, it affects the purchasing power of people's cash holdings which, in turn, affects the demand and supply of goods. This is the *real balance effect.* Patinkin denies the existence of the homogeneity postulate and the dichotomisation assumption through this effect. For this, Patinkin introduces the stock of real balances *(M/P)* held by community as an influence on their demand for goods. Thus the demand for a commodity depends upon real balance as well as relative prices. Now if the price level rises, this will reduce the real balances (purchasing power) of the people who will spend less than before. This implies a fall in the demand for goods and the consequent fall in prices and wages, The price decline increases the value of money balances held by the people which, in turn, increases the demand forgoods directly. The initial decrease in commodity demand creates a state of involuntary unemployment. Butunemployment will not last indefinitely because as wages and prices fall, the real balance effect tends to increase commodity demand directly and indirectly through the interest rate. With sufficiently large fall in wages and prices, the full employment level of output and income will be restored. Finally, even if there is the "liquidity trap", the expansion of the money supply will increase money balances, and full employment can be restored through the operation of the real balanace effect. According to Patinkin "This is the crucial point. The dynamic grouping of the absolute price-level towards its equilibrium value will—through the real balance effect—react on the commodity markets and hence on relative prices." Thus absolute prices play a crucial role not only in the money market but also in the real sector of the economy. Patinkin further points out that "once the real, and monetary data of an economy

with outside money are specified, the equilibrium values of relative prices, the rate of interest and the absolute price level are simultaneously determined by all the markets of the economy." In this way, Patinkin also introduces the real balance effect in the general equilibrium analysis.

Besides removing the classical dichotomy and the homogeneity postulate and integrating monetary and value theory through the real balance effect, Patinkin also validates the quantity theory conclusion. According to Patinkin, the real balance effect implies that people do not suffer from 'money illusion". They are interested only in the real value of their cash holdings. In other words, they hold money for 'what it will buy'. This means that a doubling of the quantity of money will lead to a doubling of the price level, but relative prices and the real balances will remain constant and the equilibrium of the economy will not be changed.

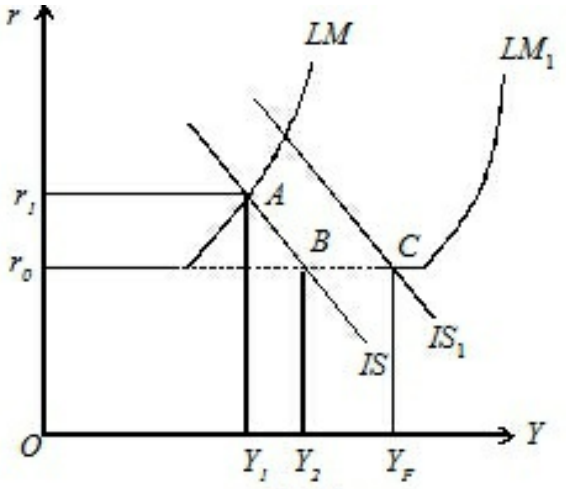


Fig. 1

The real balance effect is illustrated diagrammatically in Figure 1 by using the *IS* and *LM* technique because the *IS* curve represents the goods market and the *LM* curve the money market. To begin with, we take a situation when the economy is in equilibrium at OY_1 level of income when the *IS* and *LM* curves intersect at point *A* where the interest rate is Or_1. Assuming OY_Fas the full employment level, the pressure of unemployment is measured by $Y_1–Y_F$which causes wages and prices to fall simultaneously. This results in an increase in the real value of people's money holdings which shifts the *LM* curve to the right to LM_1. It intersects the *IS* curve at point *B* the income level OY_2 with the result that the interest rate falls to Or_0 which stimulates investment, discourages savings and increases consumption. Even when the interest rate falls to its minimum level Or_0, the level of demand in the commodity market as represented by the *IS* curve is not high enough to lead the economy to the full employment level OY_F . Rather, unemployment measured by $Y_2–$

$Y_F$remains in the economy. This much unemployment leads to a further fall in wages and prices, and to the increase in demand for consumption goods which shifts the *IS* curve to the right to IS_1 so that it intersects the LM_1curve at point *C* at the full employment level OY_F.

Thus under conditions of wage and price flexibility when the I*S* and *LM* curves shift rightward, the real balance effect ultimately leads the economy to the full employment level, even in the liquidity trap situation as shown above when investment is interest inelastic.

Conclusion. Thus the real balance effect demonstrates three theoretical points: *first,* it eliminates the classical dichotomy between value and monetary theory; *second*, it validates the conclusions of the quantity theory that in equilibrium, money is neutral and the interest rate is independent of the quantity of money through the real balance effect and *third*, the wage-price flexibility leads to full employment in the long-run and that the Keynesian underemployment equilibrium is a disequilibrium situation.

Its Criticisms

Patinkin's analysis of the real balance effect has been severely criticised by Johnson, Archibald and Lipsey, Lloyd and other economists.

1. Not Applicable in Equilibrium Situations. Johnson[2] points out that there is no need for the real balance effect so long as the real analysis is confined to equilibrium situations. The real balance is needed only to ensure the stability of the price level and not to determine the real equilibrium of the system.

2. Conceptually Inadequate. Archibald and Lipsey[3] regard Patinkin's analysis of the real balance effect as conceptually inadequate.'According to them, Patinkin traces the real balance analysis as a short-run phenomenon and does not work it out through the long-run.

---

2. Harry G. Johnson, *Essays in Monetary Theory,* 2e, 1969.

3. G.C. Archibald and R.G. Lipsey, '*Monetary* and Value Theory: A Critique of Lange and Patinkin', *R.S.E.*, XXXI (I) No. 69, October 1958.

3. Price Stability Without Real Balance Effect. Cliff Lloyd[4] has criticised Patinkin for holding the classical view that people do not suffer from 'money illusion', and that their behaviour is influenced by the real balance effect. He has shownthat the stability of the price level can be had without taking the real balance effect. According to him, byassumingthat money is available in fixed quantity and people want to hold it, will bring price stability. But 'money illusion' will not be absent.

4. Failure to Explain Increase in Monetary Wealth. Shaw[5] has criticised Patinkin for his failure to analyse the manner in which the increase in monetary wealth comes about. According to him, Patinkin simply assumes a doubling of money balances and analyses only the resultant effects. In practice, money stock does not change in this manner. "Nor, in most cases, do people experience the happy variations of helicopters carrying a surfeit of bank notes..."

Conclusion. Despite these critisms, "the introduction of the real balance effect disposes of the classical dichotomy, that is, it makes it impossible to talk about relative prices, without introducing money but it nevertheless preserves the classical proposition that the real equilibrium of the system will not be affected by the amount of money, all that will be affected will be the level of prices."

THE PIGOU EFFECT

The Pigou effect, also known as the *wealth effect*,was propounded by A.C. Pigou in 1943 to counter Keynes' argument that wage-price deflation cannot lead to automatic full employment.

Pigou fully recognised that interest-rate effect of Keynes that wage-price deflation raises investment and income through a reduction in the interest rate. But he did not agree that the real income cannot be raised to the level of full employment due to *liquidity trap*. According to Pigou, a wage-price deflation will generate automatic full employment via an increase in the level of consumption. He argued that when money wages are cut, prices

fall and the value of money rises. The rise in the value of money means a rise in the real value of assets such as stocks, shares, bank deposits, government securities, bonds, etc. For example, if prices fall by 50 per cent, the real value of each rupee will be doubled because it will purchase twice as much as it did before. The increase in the real value of fixed assets will make their owners feel richer than before. They will, therefore, save less out of their current income and spend more on consumption. This will increase aggregate demand and output, and will generate automatically full employment in the economy. As a result ofthe Pigou effect, the consumption function will shift upward (or the saving function will shift downward). In terms of the *IS* function, it means a rightward shift of the *IS* curve.

The important point in the Pigou effect is that it is based on the assumptions of flexible wage and price levels, and a constant stock of money. Therefore, it is only the *IS* curve that shifts to the right with the increase in consumption or reduction insaving when the real value of fixed assets increases. The *LM* curve is assumed to be given because of the assumption of a constant stock of money. This is because the analysis of Pigou effect runs strictly in terms of static analysis.

---

4. The Real Balance Effect: *Sine Que* What? *O.E.P.*, XIV No. 3, October 1962.

5. G.K. Shaw, *An Introduction to the Theory of Macro-economic Policy, 1*971.

Another important point is than this analysis is based upon the flexibility of absolute prices. Patinkin summarises the Pigou effect inthe following theorem: “There always exists a sufficiently low price level such that, if expected to continue indefinitely, it will generate full employment.” Algebraically, if the money supply which is assumed to be constant is *M*o and the price level is P_1 then the saving function (or consumption function) will be $S=f\ [RY\ (M_0\ /P)]$, Thus saving (*S*) depends upon the interest rate *(R)*, income (*Y*) and the ratio of given money supply to absolute prices (M_0/P). When prices fall the real value of a given stock of money rises and people reduce their saving or increase consumption thereby increasing aggregate demand. This process will automatically lead the economy to the level of full employment when decline in wages and

prices stops.In the Pigou effect, interest elasticities and positions of the saving and investment functions are irrelevant.

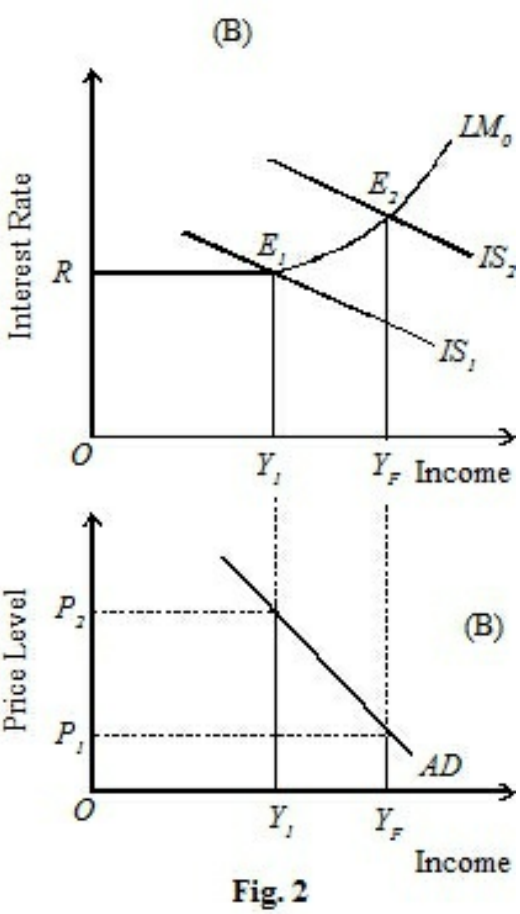


Fig. 2

The Pigou effect is illustrated in Figure 2 (A) and (B). To begin with Panel (A) of the figure, suppose theeconomy is at Y_1 level of income as determined by IS_1 and LM_0 curves at E_1. Now wage-price deflation starts which raises the consumption function such that the IS_1 curve shifts rightward to IS_2 . Given the LM_0 curve, the IS_2curve intersects the LM_0curve at E_2thereby raising the income level from OY_1 to $OY_F$the full employment level. Panel (B) of the figure shows that as the price level falls from P_2 to P_1 with reduction in money wages, income increases from OY_1 to the full employment level OY_F through the increase in aggregate demand via the Pigou effect. This is shown by the downward sloping aggregate demand curve *AD.*

Its Criticisms

The Pigou effect led to a sort of warfare among the neo-classicals and the neo-Keynesians. It appeared that the former had scored a victory over the latter by providing theoretically that if wage-price deflation is incorporated in the Keynesian model, it will automatically lead to full employment of the economy. But the neo-Keynesian resistance to the Pigou effect has been so strong that economists have pointed out a number of defects in Pigou's analysis of wage-price deflation as a means to automatic full employment. The following arguments are advanced against the Pigou analysis.

1. Ignores Distributional Effects. The Pigou effect assumes that the

depressing effect of a reduction in the price level is offset by its stimulating effect on creditors. Therefore, the private sector being a creditor to the government is stimulated by a reduction in the price level. But there are debtors other that the government. So a price decline will have different reactions on debtors and creditors. Creditors are encouraged to increase their expenditures while debtors are discouraged. But if the debtors are discouraged more than the creditors are encouraged, the price decline will not have encouraging effects on expenditures. Thus Pigou overlooked the possibility of microeconomic "distributional effects."

2. Neglects the Effect of Wage-Price Deflation on Firms. The Pigou effect considers only the effects of a change in real balances on consumption or saving of the household sector. It neglects the influence of real balances on firms. This is a serious defect in the Pigou analysis because the motivations of households and firms are different. Investment decisions of firms are favourably influenced when their real balances increase. But if firms are debtors and a price decline continues to persist, it will cause bankruptcies of firms. When debtors become bankrupt, creditors also lose. This will have a depressing effect on the economy. Thus a price decline will not lead to increase in aggregate demand and to automatic full employment in the economy.

3. Effect on Saving Uncertain. Another defect of the Pigou analysis is that it assumes definite knowledge about the effect on saving (or consumption) of a increase in the real balances. As a matter of fact, little is known about this. Moreover, only a small proportion of the lower income groups which form the majority of the population in a country, possess appreciable amount of assets. Thus very few people save and those who save do not increase their consumtion expenditures with the increase in the real value of their assets, following a price decline. Rather, they like to save more as their assets increase.

4. Wage-Price Deflation not Once-for-all. The Pigou analysis is based on the assumption of once-for-all wage-price decline. It, therefore, regards a price decline as a temporary phenomenon which is likely to be reversed when recovery starts. In practice, once-for-all price decline is not obtained. Even if it exists, it is not reasonable to suppose that people will

feel richer and increase their consumption expenditures at a single price decline.

5. Difficult to Measure Quantitatively. The Pigou effect is difficult to weigh quantitatively, that is, how much consumer expenditure will increase for any given decline in the price level. If a major deflation is required, this in itself rules out reliance on the Pigou effect as a practical means of restoring an economy to full employment. According to Shapiro, "A hyperdeflation may satisfy the purely theoretical requirements of the Pigou effect, but in practice it might also produce economic distress leading to riots and even revolution."[6]

6. Neglects Price Expectations. Another weakness of the Pigou effect is that it neglects the role of price expectations. It is not that the price level declines only once in practice. Rather, once there is a wage-price deflation, it creates expectations of a further decline in prices. This is because of the generation of pessimism among businessmen and consumers who have a tendency to reduce their expenditures in the expectation of further decline in prices. Under the circumstances, the Pigou analysis is not likely to lead to automatic full employment.

7. Pigou Effect disappears in the Lower Turning Point of a Trade Cycle. Taking the Pigou analysis in relation to the trade cycle, a once-for-all wage-price reduction will not lead to complete recovery and to full employment. Suppose the wage-price deflation has been completed and the consequent rise in real balances and in the consumption function takes the economy to the level of recovery. But as recovery continues, prices will begin to rise and the real balances with the public will start declining. So the Pigou effect will disappear once the lower turning point of a trade cycle is reached. According to Hansen, "It applies to a situation like that of 1936-40 in the United States, when prices *had* fallen and become established *at a lower level."*

---

6. Edward Shaprio, *op. cit.*

8. Static in Nature. The Pigou effect is in terms of static analysis which assumes a constant stock of money. But the real world is dynamic where

wage-price deflation with a constant stock of money will lead to full employment only after a long period, or it might even lead to a deflationary spiral with continuous unemployment.

Conclusion. Thus the Pigou effect leading to automatic full employment is unrealistic and impracticable. As he himself wrote, "The puzzles we have been considering...are academic exercises, of some slight use perhaps for clarifying thought, but with very little chance of ever being posed on the chequer board of actual life."[7]

DIFFERENCES BETWEEN PIGOU EFFECT AND REAL BALANCE EFFECT

Both the terms *"Pigou effect" and "real balance effect"* have been coined by Patinkin. But they are not the same. Rather, they are quite different. The Pigou effect is a static analysis which consists of the effect of a wage-price deflation on consumption, given the constant stock of money or what Gurley and Shaw call *"'outside money"* which includes gold, government securities and fiat paper money. It shows that when consumption increases as a result of wage-price deflation, the *IS* curve shifts to the right such that it intersects a given *LM* curve and automatic full employment is attained in the economy.

The real balance effect is amodifiedversion of the Pigou effect given by Patinkin. It is a dynamic analysis which comprises both the Pigou effect and the Keynes effect. The Keynes effect shows that as a result of wage-price deflation when the real value of money increases, smaller transactions balances are needed which create more demand for securities. Consequently, the prices of securities rise and the rate of interest falls at each level of income so that the *LM* curveshifts to the right.

Thus the operation of the Keynes effect shifts only the *LM* curve to the right and that of the Pigou effect only the *IS* curve to the right. But in the real balance effect both the *LM* and *IS* curves are shifted to the right till they intersect at the level of full employment. In the real balance effect, elasticities of the *IS and LM* curves are irrelevant. The *LM* curve may be perfectly elastic i.e., in the Keynesian liquidity trap region, or the *IS* curve maybeperfectly inelastic, the level offull employment is automatically

attained.

EXERCISES

1. Examine Patinkin's contribution to the analysis of the relationship between money supply, interest rate and price level.
2. Explain how Don Patinkin's Real Balance Effect is an improvement over the classical version of the Quantity Theory.
3. Explain the Real Balance Effect.How does it differ from the Pigou Effect and Keynes Effect?
4. Explain the Pigou Effect. In what way Patinkin has improved over it in his Real Balance Effect?
5. Examine Pigou's view that price-wage flexibility leads to full employment in the economy.

---

7. "Economic Progress in a Stable Environment," *Economics*, Vol. XIV, 1*947*.

CHAPTER 36

WAGE-PRICE FLEXIBILITY AND FULL EMPLOYMENT

INTRODUCTION

In the chapter on the Classical Theory of Employment, it was observed that the classicists advocated a reduction in money wages to achieve full employment in the economy. Keynes was against wage-cutting and argued for increase in effective demand to remove unemployment. In this chapter, we discuss in detail the classical and Keynesian views on the relationship between wages and employment.

THE CLASSICAL VIEW

The classical economists believed that there was always full employment in the economy. In case of unemployment, a general cut in money wages would take the economy to the full employment level. Their argument runs as follows:

In a competitive economy when money wages are reduced, they lead to reduction in cost of production and consequently to the lower prices of products. When prices fall, demand for products will increase and sales will be pushed up. Increased sales will necessitate the employment of more labour and ultimately full employment will be attained.

The classical view is based on the assumption that changes in money wages are related *directly and proportionately to real wages*. So when the

money wage rate is reduced, the real wage is also reduced to the same extent. Consequently, unemployment is reduced and full employment prevails.

But it is at the equality of demand and supply of labour at a particular wage rate that full employment is achieved. The demand for labour is a decreasing function of the real wage rate. If W is the money wage rate, P is the price of the product, and MP is the marginal product of labour, we have

$$W = P \times MP \qquad or$$

$$W/P = MP$$

Since MP declines as employment increases, it follows that the level of employment increases as the real wage (W/P) declines. The downward sloping demand curve for labour D is shown in Figure 1 (A). The supply of labour is assumed an increasing function of the real wage. This is shown by the upward sloping S curve in the figure. This curve indicates that the rise in real wage will induce more workers to offer themselves for employment. The interaction of S and D curves at point E shows the full employment level N_F at the real wage W/P. If the real wage is maintained at a higher level W/P_1, there is excess supply of labour by ds and N_1N_F labour is unemployed. It is only when the wage rate is reduced to W/P that unemployment disappears and the level of full employment N_F is attained.

Given the stock of capital equipment, technological know-ledge, and supplies of natural resources, a specific relation exists between the level of national income (or output) and employment. National income is an increasing function of the number of workers. The relation between national income and employment is indicated by the curve OP in Panel (B) of the figure. The level of income Y_F corresponds to the full employment level N_F of Panel (A) of Figure 1. As employment increases, the OP curve flattens out which implies that the national income increases less than in proportion to the increase in employment due to the operation of the laws of diminishing returns.

In Panel (C) of Figure 1, the equilibrium of the product and money markets of the economy is depicted in terms of the *IS and* L*M* curves. Taking the full employment level of income Y_F of Panel (B), the *IS* and *LM* (P_0) curves intersect at point *E* and the interest rate is R_F .

In the classical theory of employment, the price level is a variable. A change in the price level will affect the demand for money for transactions purposes. Hence it will change the position of the *LM* curve. Suppose the initial price level is P_0 which gives equilibrium interest rate R_F and income Y_F in Panel (C). Now the price level increases to P_1 above the full employment level P_0. It reduces the real money supply and shifts the *LM* curve to the left to *LM* (P_1). The new equilibrium point is E_1 with R_1 Y_1 combination of interest rate and income level.

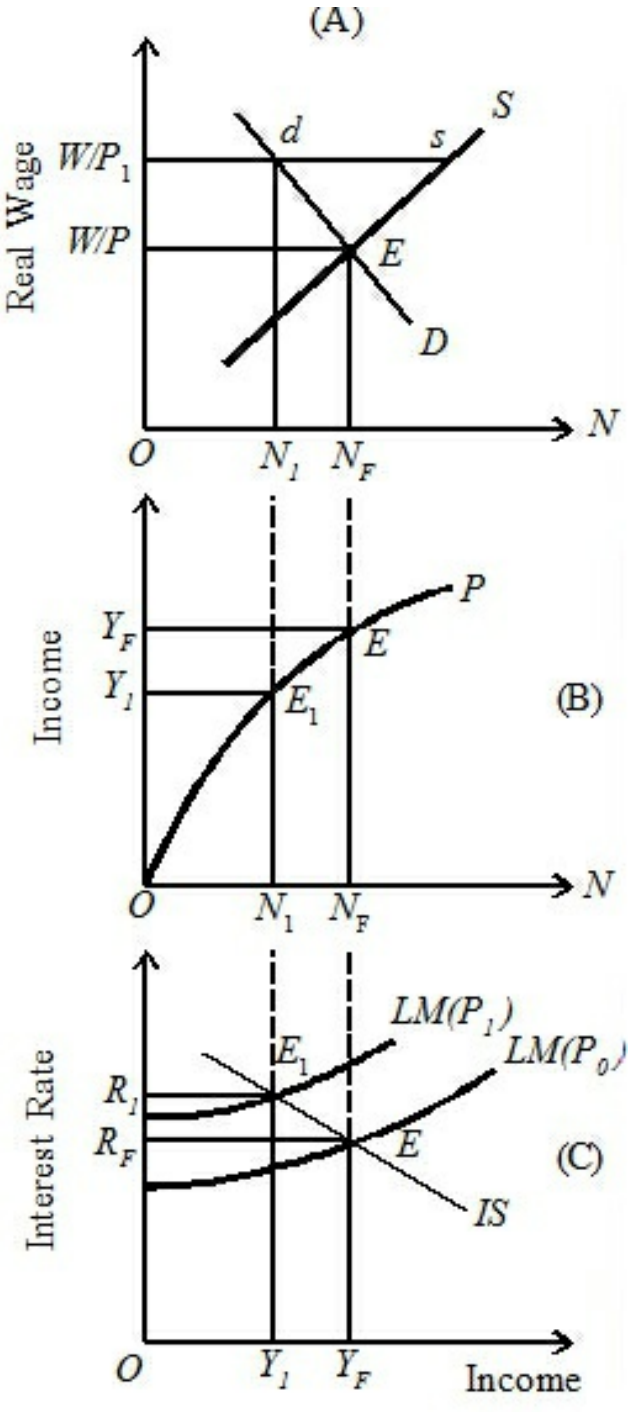


Fig. 1

With flexible prices and wages, the price level tends to settle at *Po*, thereby establishing full employment. To illustrate the mechanism by which the full employment equilibrium is attained, suppose the real wage is W/P_1 and employment N_1 in Panel (A), the level of income corresponding to this level of employment is Y_1 in Panel (B), and the price

level corresponding to this income level is P_1 , as represented by the *LM* (P_1) curve in Panel (C). At the real wage rate W/P_1 the supply of labour exceeds the demand for labour by ds. It means that N_1N_F workers are unemployed who will bid down the money wage in order to obtain employment. As the money wage falls, cost of production and prices of products fall. As wages and prices fall, the amount of money needed for transactions purposes is reduced. Thus, the supply of money exceeds the demand for money at the prevalent interest rate R_1. The excess money supply shifts the curve from *LM.* (P_1) rightward to *LM* (P_0). As a result, equilibrium is established at *E* when the economy attains the level of income Y_F and interest rate R_F in Panel (C), and employment reaches the full employment level N_F at the real wage rate *W/P* in Panel (A).

Keynes's Criticisms of the Classical View

Keynes did not accept the classical view that reduction in money wages led to full employment. He emphasised that unemployment could be removed by raising effective demand. His main objections to the classical view were as follows:

1. Not Applicable to Economy. The classical view that cut in money wages tends to increase employment is based on the partial equilibrium analysis. It is applicable to an industry. But if applied to the economy as a whole, it reduces employment rather than increasing it. According to Keynes, when money wages are reduced in the economy, they will reduce money incomes of the workers who will reduce their demand for products. Total outlay will fall and lead to a decline in effective demand and employment. Thus the main defect in the classical view was the failure to recognise the dual nature of wages as costs and incomes. The classicists only considered the cost aspect and neglected the income aspect.

2. No Need for Cut in Money Wages. Keynes did not agree with the classical view that reduction in money wages was essential to achieve full employment. According to him, full employment could be attained through monetary and fiscal policies by keeping the money wage rate constant.

3. Inverse Relation between Money and Real Wages. As against the classical view that the relation between money wages and real wages is direct and proportional. Keynes pointed out that the relation between the two is *inverse.* According to him, when money wages are reduced, real wages rise. This is because a fall in money wages will lead to a more than proportionate fall in prices. When prices fall, real wages rise because of the increase in the value of money. Thus it is not possible for the real wage to fall with a reduction in the money wage rate. Hence the entire classical argument, based on the wrong relation between money wage and real wage, falls down and it is not possible to achieve full employment by reduction in money wages.

4. Decline in Outlay, Demand and Prices. As a corollary to the above, since a cut in money wages will not lead to a reduction in real wages, Keynes believed that a reduction in money wages would lead to a proportionate decline in total outlay, demand and prices, leaving real wages intact.

5. Cut in Real Wages via Increase in Prices. Keynes agreed with the classical view that employment could be increased by a *reduction* in real wages. But the reduction in real wages was to be brought about not by a cut in money wages as the classicals believed, but by an increase in the price level. According to Keynes, an increase in aggregate demand, keeping money wages constant, would raise output, costs and prices after some time. This may happen due to three factors: (*i*) diminishing returns may operate in the short run; (*ii*) workers may demand higher wages when their demand increases due to powerful trade unions; and *(iii)* bottlenecks may appear in production in the form of non-availability of a machine, equipment, raw material, transport facilities, etc. All these factors tend to raise the cost of production and hence prices. As a result, real wages fall. Thus Keynes believed that cut in real wages increased employment but not via reduction in money wages.

6. Cut in Money Wages Impractical. Keynes further objected to a cut in money wages on practical grounds. According to him, in democratic countries where collective bargaining is premissible under the law, workers strongly resist any cut in their money wage rates.

7. Social Justice. Social justice also demands that money wages of workers *alone* should not be reduced. A cut in money wages of workers keeping the incomes of other classes as unchanged, would raise the standard of living of the latter relative to workers. Keynes, therefore, regarded "a policy of increasing the supply of money in time of depression preferable to cutting wages".

THE KEYNESIAN VIEW

Keynes started his views about the relationship between wages and employment by accepting the classical postulates that both the Law of Diminishing Returns and the Theory of Marginal Productivity operated. Since every worker is paid the wage equal to the marginal product and the law of diminishing returns operates in industry, real wages must decline for employment to increase. But this does not mean that unemployment was due to the refusal of workers to accept wages equal to their marginal product. According to .Keynes, unemployment resulted from the lack of aggregate demand. It is demand that determines employment, and employment determines the real wage rate, not the other way round.

In order to establish this, Keynes brought out the distinction between money wages and real wages. Keynes pointed out that the relation between the two is *inverse*. When money wages are reduced real wages rise. This is because a fall in money wages will lead to a more than proportionate fall in prices. When prices fall, real wages rise because of the increase in the value of money. A fall in prices will reduce aggregate demand, output, income and employment. A cut in money wages will, therefore, reduce prices and depress aggregate demand and hence employment. Like the classicists, Keynes believed that employment varies inversely with real wages. The quantity of labour demanded increases as real wages fall. But Keynes denied that a cut in money wages for the economy as a whole would necessarily lead to a cut in real wages. He believed that a reduction in money wages would lead to a proportionate decline in total outlay, demand, and prices, leaving real wages intact. To increase employment, aggregate demand should be raised while holding money wages substantially constant. Aggregate demand could be increased through monetary and fiscal policy and money wages could be

held fairly stable under modern collective bargaining practices. Thus an increase in aggregate demand, keeping money wages constant, will increase output, costs and hence prices. Consequently, real wages will fall.

Keynes argued that workers are prepared to work at the current money wage rate, even if their real wage rates are lowered by increase in prices. This is because of the existence of "money illusion" in the labour market. They will permit their real wages to be reduced by price rises without leaving the market. This is explained in Figure 2. In the figure, employment is measured on the horizontal axis and the real wage rate on the vertical axis. D_L is the demand for labour curve and S_L is the supply of labour curve. The slope of the labour supply curve is flat up to a range and then it slopes upward to the right. It signifies that in the flat segment any number of workers upto N_F level are willing to work at the prevailing money wage rate (*W)* to which they have become accustomed. So if their real wages are reduced, they will be prepared to work under the money illusion, that is, their money wages are the same even though the price level has risen. Thus up to the full employment level N_F more workers can be employed at the same money wage rate. But beyond N_F, workers would demand higher wages to supply labour. That is why the labour supply curve bends upward.

In Figure 2, at *W/P* wage rate, N_1 workers are employed when D_L and S_L curves cut at point *A.* This is the Keynesian underemployment equilibrium. At this wage level, N_1–N_F workers remain involuntarily unemployed. When the price level rises to P_1 level, the real wage rate is lowered to W/P_1 and the labour supply curve shifts downward without the workers noticing it due to money illusion. At this wage rate, the new equilibrium is established at *B* point where the demand

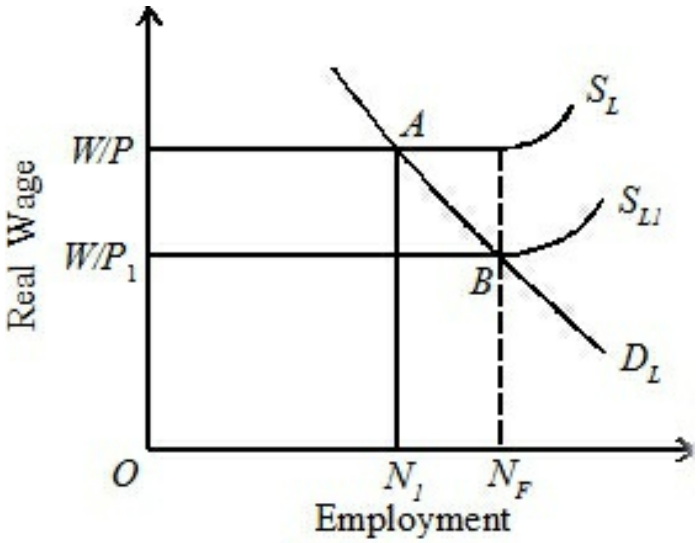


Fig. 2

for labour on the part of firms D_L equals the supply of labour S_{L1} which workers are willing to offer at the same money wage rate (*W).* Thus there is full employment of labour at N_F level.

The main difference between the classical and Keynesian propositions about cut in real wages lies in that the classicists advocated cut in real wages through cut in money wages. They believed in the flexibility of money wages. On the other hand, Keynes held that the cut in real wages would come about through the increase in aggregate demand and employment. It is by raising aggregate demand and employment first that lower real wages are followed, not the other way round. Further, Keynes believed that *money wages are not flexible, rather they are sticky downward*. This rigidity of money wages may be due to a number of institutional and other factors, such as *"(a)* powerful trade unions which are able to prevent money wages from falling, at least temporarily; *(b)* statutory provisions, such as minimum wage laws; (*c*) failure of employers to reduce wages due to a desire to retain loyal and experienced employees and to maintain morale; and *(d)* unwillingness of unemployed workers to accept reduced money wages even though they would be willing to work at lower real wages brought about by a rise in prices." Keynes, therefore, rejected a policy of flexible money wages as not only unrealistic but also because it would cause a great instability of prices, so violent perhaps as to make business calculations futile in an economic society in which we live.

Keynesian Views on Money-Wage Reductions and Employment

Despite his belief in the downward stickiness of wages, Keynes analysed the impact of flexible money wages upon the comsumption function, the *MEC* and the rate of interest, the three principal determinants of his theory. We discuss Keynes's views on money wage reductions and employment via these determinants as under.

Keynes was of the view that if reductions in money wages are not accompanied by reductions in aggregate demand, they will lead to an increase in employment. He asked, "Does a reduction in money wages have a direct tendency, *cet par*., to increase employment, "cet. par," being

taken to mean that the propensity to consume, the schedule of marginal efficiency of capital and the rate of interest are the same as before ?" But these variables are bound to change and adversely affect aggregate demand when money wages are reduced. Keynes, therefore, discussed the possible effects of a general money wage cut on aggregate demand and employment through these three variables.

Wage Cut and Propensity to Consume. According to Keynes, the effect of a general cut in money wages on the propensity to consume will be unfavourable. As a result of wage reductions, incomes of the wage-earning classes will fall in relation to that of non-wage earners and there will be some redistribution of income in favour of the latter. This will decrease total consumption since the incomes of the wage earners having a high propensity to consume fall. The rich, on the other hand, have a high propensity to save and a low propensity to consume. This causes a decline in aggregate demand, income and employment. This is called *Keynes's redistributive effect.*

Pigou, however, considered the possibility when money wage reductions may raise the propensity to consume. He argued that as a consequence of general wage cut, prices will fall and the value of money will rise. With the rise in the value of money, the real value of money assets such as stocks, shares, bank deposits, government bonds and securities, etc. will rise. The owners of these fixed assets will feel themselves richer than before and their propensity to consume will be strengthened and the propensity to save weakened. This is known as the *Pigou Effect* which will tend to increase aggregate demand and employment.

The Pigou effect fails to increase the propensity to consume much due to three reasons. *First*, very few people possess money assets. So the influence on the overall propensity to consume is very insignificant. *Second*, it is not necessary that asset-holders may at once increase their propensity to consume with the rise in the real value of their money assets. Rather, their desire to save may increase. *Third*, the stimulating effect of larger real cash balances on consumption may be offset by the discouraging effect of increased real debt burden on consumption due to lower prices, and thus leave the net effect very small.

To take Keynes's redistributive effect and the Pigou effect together, we find that Keynes's redistributive effect is stronger than Pigou's effect and the overall effect is a fall in the propensity to consume. This is due to the fact that the redistributive effect tends to reduce mass consumption whereas the Pigou effect raises the consumption of only the asset holders whose number is very small. Thus the net result of a general wage cut is to reduce the propensity to consume, aggregate demand, income and employment.

Wage Cut and the MEC. A general wage cut is bound to influence the *MEC,* investment and employment through business expectations. If a wage cut leads to expectations of further reductions in wages, the *MEC* will fall because entrepreneurs will postpone their investment plans till they are sure that wages have fallen to the lowest level. This will have an unfavourable effect on employment. But if they do not expect wages to fall further or expect them to be higher in future, the *MEC* will be favourably affected. Entrepreneurs will plan to invest more and thus employment will increase.

But wage reductions may adversely affect the *MEC* if trade unions oppose cut in money wages by resorting to strikes. It will lead to unrest and a fall in business expectations and the *MEC.* No entrepreneur will be prepared to invest under such conditions. Labour psychology thus plays an important part in offsetting otherwise favourable expectations.

Moreover, a fall in prices due to wage reductions will increase the real burden of private debt and adversely affect business expectations, *MEC* and investment. Similarly with falling prices, the real burden of public debt will also increase and the government may levy new taxation to repay the debt and service charges. This will again have a depressing effect on investment and employment.

In an open system, when prices of export goods fall due to a wage cut, exports will increase and a rise in entrepreneurs' expectations will increase investment. But this is only possible if other countries do not cut wages.

We may conclude that "taking account of the actual practices and institutions of the contemporary world, a stable wage policy is likely to

have more favourable effect on business expectations than a flexible policy under which wages would drift downwad by easy stages as unemployment increased."[1]

Wage Cut and the Interest Rate. Reductions in money wage rate will have a favourable effect on the level of investment and employment by lowering the rate of interest. When prices fall due to a general wage cut, the demand for money for transactions and precautionary motives will fall and for speculative motive increase. Thus more money will be available for speculative motive and the rate of interest would fall. If the investment demand curve is interest elastic, a lower interest rate would tend to increase the level of investment and employment. This process of increase in investment and employment via reduction in wages and rate of interest is called the “Keynes Effect.”

The extent to which wage cuts will reduce the rate of interest and affect favourably investment and employment depends upon the amount of wage reductions. If wage reductions are moderate, they may prove inadequate in lowering the rate of interest so as to encourage investment. On the other hand, an immoderate reduction in wages might create political and social unrest through trade union opposition, cause unfavourable business expectations and bring a fall in investment and a rise in liquidity preference. Thus the effect of a wage cut on investment and employment via the interest rate does not lead to any definite conclusion.

---

1. A.H. Hansen, *A Guide to Keynes*, p.177

Keynes, therefore, favoured a flexible money policy to a flexible wage policy to increase investment and employment. The rate of interest can be lowered by an expansion in money supply by the monetary authority, keeping money wages intact. A flexible wage policy in the form of a general wage cut has certain practical difficulties. Such a policy would be vehemently opposed by trade unions and cause a great instability of prices, so violent perhaps as to make business calculations futile. It will make the maintenance of continuous full employment impossible. Moreover, it will tend to increase the real burden of private and public debt with its adverse effects on investment. The same results can,

however, be achieved by a flexible money policy without the adverse effects of a flexible wage policy. Keynes was thus justified in saying that "only a foolish person would prefer a flexible wage policy to a flexible money policy."

THE KEYNES EFFECT*

The process of increase in investment and employment via reduction in money wages and interest rate is called the Keynes effect. Keynes maintained that as money wages and prices fall, money incomes also fall which means increase in real output and incomes. People will, therefore, need less money for transactions and precautionary purposes. Assuming no change in the supply of money, this means that people have more money in liquid form than they need. It will increase the demand for money for speculative purposes. They will, therefore, invest surplus money in purchasing securities and bonds. The increase in the demand for such financial assets will raise their prices which will reduce the rate of interest. Falling interest rate, in turn, will raise investment, national income and employment.

But Keynes did not believe that a policy of flexible wages and prices can lead to full employment. It can only lead to underemployment equilibrium. Keynes suggested two factors that may prevent the economy from reaching the level of full employment via wage-price flexibility and thus break down the Keynes effect. *First*, in the beginning with the reduction in money wages and prices, the demand for money may fall and lead to falling interest rate. But when the interest rate reaches a very low level, the liquidity trap may operate. *Second*, if investment is quite insensitive (inelastic) to a fall in the rate of interest. In other words, if the slope of the *MEl* curve is very steep, even a substantial reduction in the rate of interest will have little effect on investment.

Further, the extent to which wage cuts will reduce the interest rate and increase investment and employment depends upon the amount of wage reductions. If wage reductions are moderate, they may be inadequate in lowering the rate of interest. On the other hand, an immoderate reduction in wages might create social unrest through trade union opposition and

cause unfavourable business expectations, thereby bring about a fall in investment. Keynes concluded thus: "There is, therefore, no ground for the belief that a flexible wage policy is capable of maintaining a state of continuous full employment...The economic system cannot be made self-adjusting along these lines."

---

* Student should also study the Pigou Effect.

The Keynes effect is illustrated in Figure 3 (A) and (B). The original equilibrium level in the economy is depicted by the intersection of *IS* and *LM* curves at point *E* with *R=Y* combination of interest rate and income level, and P_3 price level. Consider a position when the money wage rate falls and the price level is reduced to P_3 in Panel (B) of the figure. This increases the real supply of money and shifts the *LM* curve in Panel (A) to the right to LM_1 where it intersects the *IS* curve at point E_1. This point is associated with the lower interest rate R_1 and higher income level Y_1. A further fall in the price level to P_2 shifts the *LM* curve from LM_1 to LM_2 thereby creating a new equilibrium point with the *IS* curve at E_2. This reduces the interest rate to R_2 and raises income to Y_2. But a further fall in money wage rate and the price level will not reduce the interest rate below R_2 because the *IS* curve cuts the LM_2 curve in the liquidity trap. Further reduction in the price level

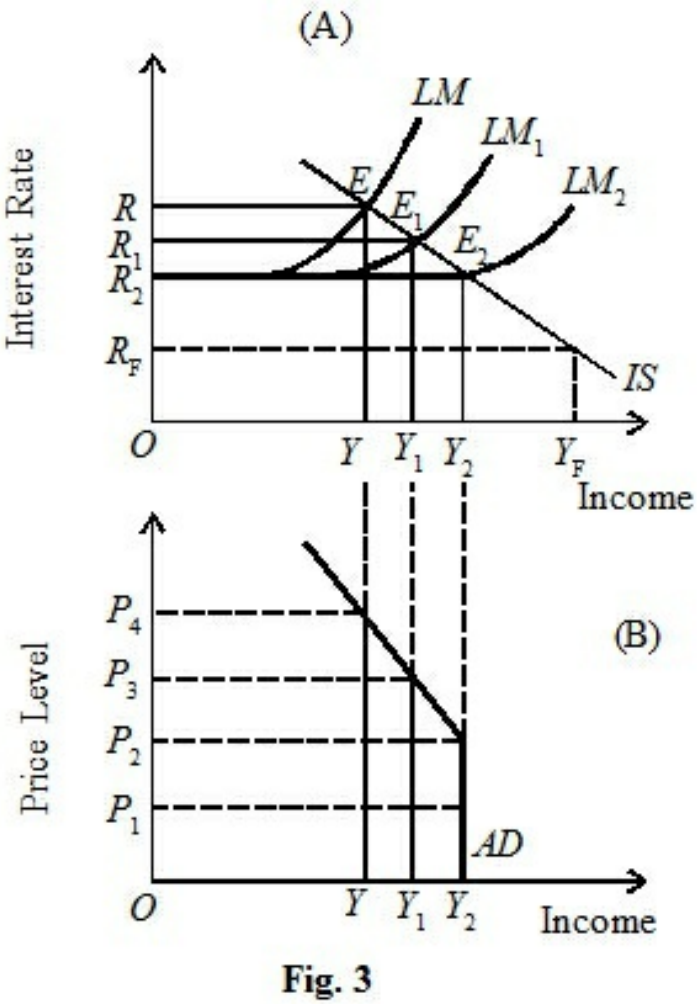


Fig. 3

will continue to release money from transactions balances, but this money will be absorbed in speculative balances without further reducing the interest rate. Thus no amount of wage-price reduction will be able to raise income to the full employment level Y_F by lowering the interest rate to R_F level, as shown in Panel (A) of the figure. The figure in the lower Panel (B) shows that with every fall in the price level, the real value of money supply rises which lowers the interest rate and raises the level of investment and income. So the aggregate demand curve *AD* is negatively sloped upto the price level P_2. But in the liquidity trap, a further fall in the money wage rate leads to a proportionate fall in the price level so that there is no change in investment, income and employment. This is depicted in Panel (B) of the figure by the vertical position of the *AD* curve when the price level falls from P_2 to P_1 and the level of income remains constant at Y_2.

Its Criticisms

The Keynes effect has been criticised by Hansen, Patinkin and other economists on the following grounds:

1. Ignores other Factors. Patinkin has criticised Keynes for confining the effects of money wage reduction only in the transactions and precautionary cash balances and making the speculative demand for money depend solely on the interest rate. This has led to three weaknesses in this analysis of wage-price flexibility.

(a) Real Balance Effect. Keynes overlooked the real-balance effect on consumption and on the demand for money. According to Hansen, "There is no evidence that Keynes ever thought of the Pigou effect. It had been stated only vaguely in the long controversy *(prior to the appearance of the General Theory)* about the consequences of declining prices." Patinkin observes in this connection : "It seems likely that Keynes did recognise the influence of wealth on consumption (or rather savings) but thought of this influence only in terms of non-monetary assets."

(b) Money Illusion. Keynes assumed “money illusion” in the speculative demand for money. Accoding to Patinkin, “The Keynesian monetary

theory never permits the speculative demand for money to absorb increased supply of money except at a lower interest rate. When wage-price deflation leads to increase in the supply of money, the fall in the rate of interest builds a money illusion into the speculative demand for money."

(c) Price Level Expectations. Keynes neglected the influence of price-level expectations on the demand for money. His analysis is based on a once-and-for-all reduction in money wages and prices which does not give rise to any expectations of further reductions in wages and prices. In practice, one wage-price reduction often leads to further reductions and such pessimistic expectations will adversely affect investment decisions.

2. Liquidity Trap an Extreme Case. Keynes's notion of the liquidity trap is an extreme case, because that could occur only during a deep depression. Since there has been no depression of the type of the Great depression, economists no longer worry about the liquidity trap and they even doubt its existence. However, some followers of Keynes hold the view that the liquidity trap exists at a zero interest rate because no lender would lend at a negative interest rate.

3. Investment not Insensitive to Interest Rate. Patinkin does not agree with Keynes that wage-price flexibility will not lead to full employment if investment is completely insensitive to a fall in the interest rate. According to him, even if the saving and investment functions are completely insensitive to change in the interest rate, a wage reduction will still be stimulatory through its effect on real balances and will lead to full employment. Moreover, it is not possible to predict the net result of these redistribution effects on aggregate investment or consumption expenditure. Creditors might increase their expenditure due to gains and debtors might reduce it due to losses. So the net effect on aggregate expenditure of investment or consumption is unpredictable.

4. Adverse Effects of Wage-Price Deflation. The policy of wage-price deflation is likely to lead to serious distributional effects on debtors and creditors unless the terms of indebtedness are not changed in conformity with lower wages and prices. When money wages and prices are reduced, the real value of monetary debt will increase. As a result, creditors will

become rich and debtors will become poor to the extent of the debt.

5. Wage-Price Deflation a Theoretical Possibility. Keynes himself regarded the policy of wage-price deflation for expanding employment as only a theoretical possibility. For, there are many institutional barriers to wage reductions. In reality, wage cuts reduce money incomes and expenditures of wage earners. They lead to reduction in output, employment and income, rather the other way round.

6. Wage-Price Deflationary Process Absurd. According to Hansen, the wage-price deflationary process to raise employment is patently absurd. If what is wanted is more money and more liquid assets relative to income, the obvious method to achieve this result is through monetary policy. Keynes himself asserted that a flexible wage-price policy was incapable of maintaining continuous full employment. He, therefore, favoured a flexible monetary policy to a flexible wage policy.

FLEXIBLE WAGE POLICY VS FLEXIBLE MONETARY POLICY

Keynes did not agree with the classical economists that full employment could be achieved by a flexible wage policy. He favoured a flexible monetary policy to a flexible wage policy due to certain practical difficulties involved in following the latter policy. In order to discuss Keynses's views on the issue, it is instructive to study briefly the classical view.

The classical economists believed that there was always full employment in the economy. Unemployment was due to wage rigidity. So a general cut in money wages would move the economy to the full employment level. Their argument is that in a competitive economy a money wage cut will reduce cost of production and prices. A fall in prices will increase demand and sales of products, thereby necessitating the employment of more labour. This process will ultimately lead the economy to the full employment level. The classical argument is based on the assumption that a change in money wage rate is related *directly and proportionately* to real wages. So when money wages are reduced, real wages are reduced to the same extent.

Keynes did not agree with the classicists that a cut in money wage rate is essential to attain full employment. He objected to the policy of downward flexibility of money wages on the following grounds.

According to Keynes, a general cut in money wages reduces employment rather than increasing it. A cut in money wages reduces the money incomes of workers who reduce their demand for products. This will lead to decline in aggregate demand, output and employment.

Further, when wages of workers alone are reduced, their income falls in relation to that of non-wage earners. As the propensity to consume of the workers is high and that of non-wage earners low, the aggregate consumption will decline. This will cause a decline in aggregate demand, output and employment.

Again, a gereral cut in money wages is bound to adversely affect the *MEC*, investment and employment in the economy. If businessmen expect wages to fall further with a wage cut, the *MEC* will fall and businessmen will postpone their investment plans. This will adversely affect employment. Moreover, if trade unions oppose cut in the money wage rate by resorting to strikes, the *MEC* will again fall. This is because labour unrest will lead to fall in business expectations and the *MEC*. Businessmen will be hesitant to invest under such conditions, thereby adversely affecting employment.

Keynes did not favour a flexible wage policy in the form of a cut in money wages on certain practical grounds also. According to him, such a policy would be opposed by trade unions. They would not accept an all-round reduction in money wages. In democratic countries, where collective bargaining is permissible, reduction in money wages is an impracticable possibility. It is, therefore, much easier to follow a flexible monetary policy than a flexible wage policy.

Further, social justice requires that money wages of worker *alone* should not be cut. A cut in money wages of workers, while keeping the incomes of other sections of the community as unchanged, would lower the standard of living of the former relative to the latter. Social justice demands that all classes of society should be treated equally and workers

alone should not suffer due to cut in their money wages. That is why Keynes preferred a policy of increasing the money supply in depression to a policy of cut in money wages because it would affect the real incomes of all sections as prices rise.

Last but no the least, when prices fall due to a cut in money wages, they increase the real burden of debt on debtors. Since the majority of debtors are businessmen, such a policy will adversely affect business expactations, *MEC*, investment, output and employment. On the other hand, a policy of increasing the supply of money would reduce the burden of debt because it would increase the price level.

The so-called Keynes effect which Keynes put forth in favour of a policy of cutting money wages was simply meant to analyse the possibility of increasing investment and employment via reduction in wages and interest rate. But he did not think that a policy of cut in money wage rates with an unchanged money supply was likely to be of much practical importance because it was incapable of achieving full employment due to the liquidity trap. According to Keynes, "The economic system cannot be made self-adjusting along these lines." Moreover, a flexible wage policy involves high social costs and is impracticable. He, therefore, pointed out that the same result could be obtained from a rise in the money supply and unchanged money wage rates. In other words, a policy of increasing the supply of money can lead to increase in employment by lowering the interest rate, like the policy of a cut in money wages. Such a policy is not only realistic but also involves far less social costs. Thus Keynes favoured a flexible monetary policy of increasing the money supply in order to raise employment, keeping money wages unchanged. Keynes wrote in this connection : "Having regard to human nature and our institutions, it can only be a foolish person who would prefer a flexible wage policy to a flexible money policy, unless he can point to advantages from the former which are not obtainable from the latter."

SELECTED READINGS

J.M. Keynes, *The General Theory of Employment, Interest and Money*, 1936,

A.H. Hansen, *A Guide to Keynes,* 1953, Ch. 10.

Edward Shapiro, *Macroeconomic Analysis,* 4/e, 1978, Ch. 16.

T.F. Dernburg and D.M. McDougall, *Macroeconomics,* 5/e, 1976.

Don Patinkin, *Money, Interest and Prices,* 2/e, 1965.

EXERCISES

1. Trace the effects of a cut in money wages on level of employment, bringing out in your answer the view points of classical economists and the Keynesians.
2. Compare the classical and Keynesian positions in regard to the relationship between wage flexibility and full employment.
3. 'Unemployment equilibrium can occur only if money wages are inflexible downward'. Discuss critically this assessment of Keynes's contribution to the theory of employment.
4. Discuss critically the Keynes Effect.
5. "Only a foolish person would prefer a flexible wage policy to a flexible money policy." Discuss this statement.

PART-V

INFLATION AND BUSINESS CYCLES

CHAPTER 37

INFLATION AND DEFLATION

INTRODUCTION

Inflation is a highly controversial term which has undergone modification since it was first defined by the neo-classical economists. They meant by it a galloping rise in prices as a result of the excessive increase in the quantity of money. They regarded inflation "as a destroying disease born out of lack of monetary control whose results undermined the rules of business, creating havoc in markets and financial ruin of even the prudent."[1]

But Keynes in his *General Theory* allayed all such fears. He did not believe like the neo-classicists that there was always full employment in the economy which resulted in hyper-inflation with increases in the quantity of money. According to him, there being underemployment in the economy, an increase in the money supply leads to increase in aggregate demand, output and employment. Starting from a depression, as the money supply increases, output at first rises proportionately. But as aggregate demand, output and employment rise further, diminishing returns start and certain bottlenecks appear and prices start rising. This process continues till the full employment level is reached. The rise in the price level during this period is known as *bottleneck inflation* or "semi-inflation". If the money supply increases beyond the full employment level, output ceases to rise and prices rise in proportion with the money supply. This is true inflation, according to Keynes.

---

1. R.J. Ball and Peter Doyle (ed.) *Inflation, 1967,* p. 7

Keynes's analysis is subjected to two main drawbacks. *First*, it lays emphasis on demand as the cause of inflation, and neglects the cost side of inflation. *Second*, it ignores the possibility that a price rise may lead to further increase in aggregate demand which may, in trun, lead to further rise in prices.

However, the types of inflation during the Second World War, in the immediate post-war period, till the middle of the 1950s were on the Keynesian model based on his theory of excess demand. "In the latter 1950s, in the United States, unemployment was higher than it had been in the immediate post-war period, and yet prices still seemed to be rising, at the same time, the war time fears of postwar recession had belatedly been replaced by serious concern about the problem of inflation. The result was a prolonged debate...On the one side of the debate was the 'cost-push' school of thought, which maintained that there was no excess demand...On the other side was the "demand-pull" school...Later, in the United States, there developed a third school of thought, associated with the name of Charles Schultz, which advanced the sectoral 'demand-shift theory' of inflation...While the debate over cost-push versus demand-pull was raging in the United States, a new and very interesting approach to the problem of inflation and anit-inflationary policy was developed by A.W. Phillips."[2]

In the present chapter, we shall study all theories mentiond here, besides Keynes's thoery of the inflationary gap. But before we analyse them, it is instructive to know about the meaning of inflation.

MEANING OF INFLATION

To the neo-classicals and their followers at the University of Chicago, inflation is fundamentally a monetary phenomenon. In the words of Friedman, "*Inflation is always and everywhere a monetary phenomenon...and can be produced only by a more rapid increase in the quantity of money than output.*"[3] But economists do not agree that money supply alone is the cause of inflation. As pointed out by Hicks, "Our present troubles are not of a monetary character." Economists, therefore, define inflation in terms of a continuous rise in prices. Johnson defines

"inflation as a sustained rise"[4] in prices. Brooman defines it as "a continuing increase in the general price level."[5] Shapiro also defines inflation in a similar vein "*as a persistent and appreciable rise in the general level of prices*."[6] Dernberg and McDougall are more explicit when they write that "the term usually refers to a continuing rise in prices as measured by an index such as the consumer price index (CPI) or by the implicit price deflator for gross national product."[7]

However, it is essential to understand that a sustained rise in prices may be of various magnitudes. Accordingly, different names have been given to inflation depending upon the rate of rise in prices.

---

2. Harry G. Johnson, *Essays in Monetary Economics*.

3. *The Counter Revolution in Monetary Theory*, 1970.

4. Harry W. Johnson *op. cit*, p. 104.

5. F. S. Brooman, *op. cit*., p. 285.

6. Edward Shapiro, *op. cit*., p. 142.

7. *Macroeconomics*,, pp. 288.

1. Creeping Inflation. When the rise in prices is very slow like that of a snail or creeper, it is called creeping inflation. In terms of speed, a sustained rise in prices of annual increase of less than 3 per cent per annum is characterised as creeping inflation. Such an increase in prices is regarded safe and essential for economic growth.

2.Walking or Trotting Inflation. When prices rise moderately and the annual inflation rate is a single digit. In other words, the rate of rise in prices is in the intermediate range of 3 to 7 per cent annum or less than 10 per cent. Inflation at this rate is a warning signal for the government to control it before it turns into running inflation.

3. Running Inflation. When prices rise rapidly like the running of a horse at a rate of speed of 10 to 20 per cent per annum, it is called running inflation. Such an Inflation affects the poor and middle classes adversely. Its control requires strong monetary and fiscal measures, otherwise it leads

to hyperinflation.

4. Hyperinflation. When prices rise very fast at double or triple digit rates from more than 20 to 100 per cent per annum or more, it is usually called *runaway* or *galloping* inflation. It is also characterised as hyperinflation by certain economists. In reality, hyperinflation is a situation when the rate of inflation becomes immeasurable and absolutely uncontrollable. Prices rise many times every day. Such a situation brings a total collapse of monetary system because of the continuous fall in the purchasing power of money.

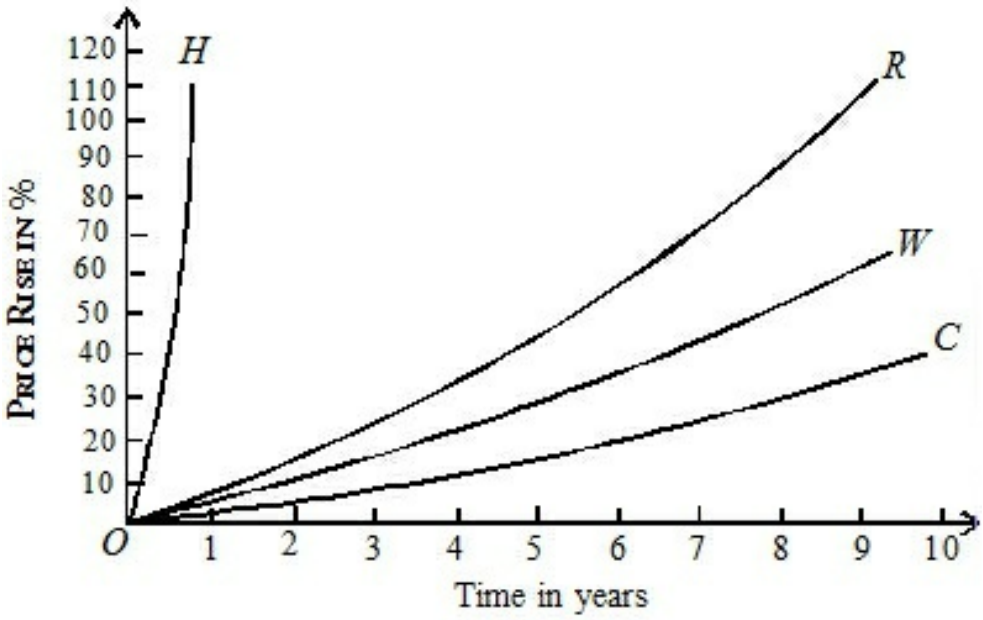


Fig. 1.

The speed with which prices tend to rise is illustrated in Figure 1. The curve *C* shows creeping inflation when within a period of ten years the price level has been shown to have risen by about 30 per cent. The curve W depicts walking inflation when the price rose by more than 50 per cent during ten years. The curve *R* illustrates running inflation showing a rise of about 100 per cent in ten years. The steep curve *H* shows the path of hpyerinflation when prices rose by more than 120 per cent in less than one year.

THE INFLATIONARY GAP

In his pamphlet *How to Pay for the War* published in 1940, Keynes explained the concept of the inflationary gap. It differs from his views on

inflation given in the *General Theory*. In the *General Theory*, he started with underemployment equilibrium. But in *How to Pay for the War*, he began with a situation of full employment in the economy. He defind an inflationary gap as an excess of planned expenditure over the available output at pre-inflation or base prices. According to Lipsey, "The inflationary gap is the amount by which aggregate expenditure would exceed aggregate output at the full employment level of income." The classical economists explained inflation as mainly due to increase in the quantity of money, given the level of full employment. Keynes, on the other hand, ascribed it to the excess of expenditure over income at the full employment level. The larger the aggregate expenditure, the larger the gap and the more rapid the inflation. Given a constant average propensity to save, rising money incomes at full employment level would lead to an excess of demand over supply and to a consequent inflationary gap. Thus Keynes used the concept of the inflationary gap to show the main determinants that cause an inflationary rise of prices.

The inflationary gap is explained with the help of the following example:

Suppose the gross national product at pre-inflation prices is Rs. 200 crores.Of this Rs. 80 crores is spent by the Government. Thus Rs. 120 (Rs. 200-80) crores worth of output is available to the public for consumption at pre-inflation prices. But the gross national income at current prices at full employment level is Rs. 250 crores. Suppose the government taxes away Rs. 60 crores, leaving Rs. 190 crores as disposable income. Thus Rs. 190 crores is the amount to be spent on the available output worth Rs. 120 crores. thereby creating an inflationary gap of Rs. 70 crores.

This inflationary gap model is illustrated as under :

1.Gross National Income at current prices	Rs. 250 Cr.	=
2.Taxes	= Rs. 60 Cr.	
3.Disposable Income	= Rs. 190 Cr.	
4.GNP at pre-inflation prices	= Rs. 200 Cr.	
5.Government expenditure	= Rs. 80 Cr.	
6.Output available for consumption at		

pre-inflation prices	=	Rs. 120 Cr.
Inflationary gap (Item 3-6)	=	Rs. 70 Cr.

In reality, the entire disposable income of Rs. 190 Crores is not spent and a part of it is saved. If, say, 20 per cent (Rs. 38 croes) of it is saved, then Rs. 152 crores (Rs. 190-Rs. 38 crores) would be left to create demand for goods worth Rs. 120 crores. Thus the actual inflationary gap would be Rs. 32 (Rs. 152-120) crores instead of Rs. 70 crores.

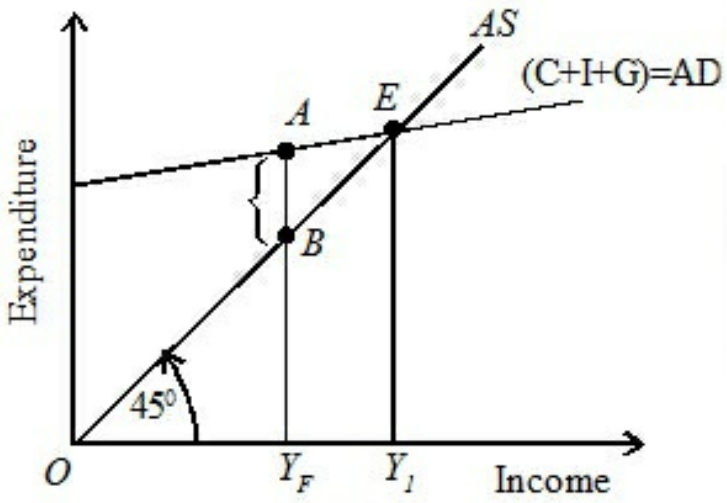


Fig. 2.

The inflationary gap is shown diagrammatically in Figure 2 where Y_F is the full employment level of income, 45° line represents aggregate supply *AS* and *C+1+G* line the desired level of consumption, investment and government expenditure (or aggregate demand curve). The economy's aggregate demand curve (*C+1+G*) = *AD* intersects the 45 per cent line (A_S) at point *E* at the income level OY_1 which is greater than the full employment income level OY_F. The amount by which aggregate demand (Y_FA) exceeds the aggregate supply (Y_FB) at the full employment income level is the inflationary gap. This is *AB* in the figure. The excess volume of total pending when resources are fully employed creates inflationary pressures. Thus the inflationary gap leads to inflationary pressures. in the economy which are the result of excess aggregate demand.

How can the inflationary gap be wiped out ?

The inflationary gap can be wiped out by increase in savings so that the aggregate demand is reduced. But this may lead to deflationary tendencies.

Another solution is a raise the value of available output to match the disposable income. As aggregate demand increases, businessmen hire more labour to expand output. But there being full employment at the current money wage, they offer higher money wages to induce more workers to work for them. As there is already full employment, the increase in money wages leads to proportionate rise in prices. Moreover, output cannot be increased during the short run because factors are already fully employed. So the inflationary gap can he closed by increasing taxes and reducing expenditure. Monetary policy can also be used to decrease the money stock. But Keynes was not in favour of monetary measures to control inflationary pressures within the economy.

Its Criticisms

The concept of inflationary gap has been criticised by Friedman[8], Koopmans[9], Salant[10], and other economists.

1. The analysis of inflationary gap is based on the assumption that full employment prices are flexible upward. In other words, they respond to excess demand in the market for goods. It also assumes that money wages are sticky when prices are rising., but the share of profits in GNP increases. So this concept is related to excess-demand inflation in which there is profit inflation. This has led to the mixing up of demand and cost inflations.

2. Bent Hansen criticises Keynes for confining the inflationary gap to the goods market only and neglecting the role of the factor market. According to him, an inflationary gap is the result of excess demand in the goods market as well as in the factor market.[11]

3. The inflationary gap is a static analysis. But the inflationary phenomena are dynamic. To make them dynamic, Keynes himself suggested the intorduction of time lags concerning receipts and expenditures of income. Koopmans has developed relationships between eggs and the rate of price increase per unit of time. He has shown with the help of spending lags and wage-adjustment lags that the speed of inflation becomes smaller, that is the inflationary gap is narrowed.

4. Holzman has criticised Keynes for applying the multiplier technique to a full employment situation. According to him, the multiplier technique is not adequate in periods of full employment and inflation. It abstracts from changes in the distribution of income. In a full employment situation, the share of one group in the national output can only be increased at the expense of another.[12]

5. Another weakness of the inflationary-gap analysis is that it is related to flow concepts, such as current income, expenditure, consumption, and saving. In fact, the increase in prices at the full employment level is not confined to prices of current goods alone. But they also affect the prices of goods already produced. Further, the disposable income which is the difference between current income and taxes, may include idle balances from the income of previous periods.

---

8. M. Friedman, "Discussion of the Inflation of the Gap", *A.E.R.*, Vol. 32, 1942.

9. T. Koopmans, "The Dynamics of Inflation", *R.E.S.*, Vol. 24, 1942.

10. W. Salant, "The Inflationary Gap, Meaning and Significance for Policy Making", *A.E.R.* Vol. 32, 1942.

11. Bent Hansen, A Study in the Theory of Inflation, 1951.

Its Importance

Despite these criticisms the concept of inflationary gap has proved to be of much importance in explaining rising prices at full employment level and policy measures in controlling inflation.

It tells that the rise in prices, once the level of full employment is attained, is due to excess demand generated by increased expenditures. But the output cannot be increased because all resources are fully employed in the economy. This leads to inflation. The larger the expenditure, the larger the gap and more rapid the inflation.

As a policy measure, it suggests reduction in aggregate demand to control inflation. For this, the best course is to have a surplus budget by raising taxes. It also favours saving incentives to reduce consumption

expenditure. “The analysis of the inflationary gap in terms of such aggregate as national income, investment outlays and consumption expenditures clearly reveals what determines public policy with respect to taxes, public expenditures, savings campaigns, credit control, wage adjustment—in short, all the conceivable anti-inflationary measures affecting the propensities to consume, to save and to invest which together determine the general price level.”[13]

DEMAND-PULL OR MONETARY THEORIES OF INFLATION

Demand-pull inflation or excess demand inflation is the traditional and most common type of inflation. It takes place when aggregate demand is rising while the available supply of goods is becoming less. Goods may be in short supply either because resources are fully utilised or production cannot be increased rapidly to meet the increasing demand. As a result, prices begin to rise in response to a situation often decribed as “too much money chasing too few goods.”

There are two principal theories about the demand-pull inflation that of the monetarists and the Keynesians. We shall also discuss a third one propounded by the Danish economist, Bent Hansen.*

1. Monetarist View or Monetary Theory of Inflation

The monetarists emphasise the role of money as the principal cause of demand-pull inflation. They contend that inflation is always a monetary phenomenon. Its earliest explanation is to be found in the simple quantity theory of money. The monetarists employ the familiar indentity of Fisher’s Equation of Exchange.

$$MV = PQ$$

where M is the money supply, V is the velocity of money, P is the price level, and Q is the level of real output.

Assuming V and Q as constant, the price level (P) varies proportionately with the supply of money (M). With flexible wages, the economy was

believed to operate at full employment level. The labour force, the capital stock, and technology also changed only slowly over time. Consequently, the amount of money spent did not affect the level of real output so that a doubling of the quantity of money would result simply in doubling the price level. Until prices had risen by this proportion, individuals and firms would have excess cash which they would spend, leading to rise in prices. So inflation proceeds at the same rate at which the money supply expands. In this analysis the aggregate supply is assumed to be fixed and there is always full employment in the economy. Naturally, when the money supply increases it creates more demand for goods but the supply of goods cannot be increased due to the full employment of resources. This leads to rise in prices. But it is a continuous and prolonged rise in the money supply that will lead to true inflation.

---

12. F. D. Holzman, "Income Determination in Open Inflation", *R.E.S.*, Vol.32, 1950.

13. K.K. Kurihara, *Monetary Theory and Public Policy*, p. 45.

* B. Hansen's Monetary Theory of Inflation is not to be given unless asked for.

This classical theory of inflation is explained in Fig. 3 where the quantity of money is taken on horizontal line and the price level on vertical line. When the quantity of money is *OM*, the price level is *OP*. When the quantity of money is doubled to OM_2 the price level is also doubled to P_2. Further, when the quantity of money is increased four-fold to M_4, the price level also increases by four times to P_4. This relationship is expressed by the curve $P = f(M)$ from the origin at 45°.

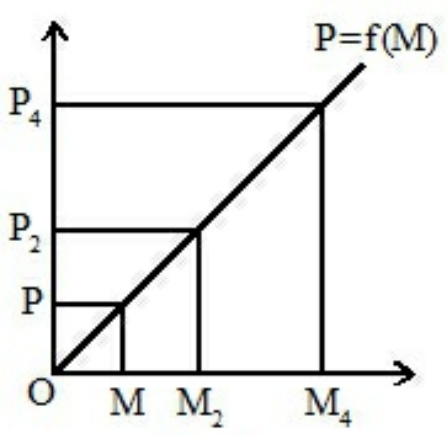


Fig. 3.

Friedman's View

Modern quantity theorists led by Friedman hold that "inflation is always and everywhere a monetary phenomenon that arises from a more rapid expansion in the quantity of money than in total output." He argues that

changes in the quantity of money will work through to cause changes in nominal income. Inflation everywhere is based on an increased demand for goods and services as people try to spend their cash balances. Since the demand for money is fairly stable, this excess spending is the outcome of a rise in the nominal quantity of money supplied to the economy. So inflation is always a monetary phenomenon.

Next Friedman discusses whether an increase in money supply will go first into output or prices. Initially, when there is monetary expansion, the nominal income of the people increases. Its immediate effect will be to increase the demand for labour. Workers will settle for higher wages. Input costs and prices will rise. Profit margins will be reduced and the prices of products will increase. In the begining, people do not expect prices to continue rising. They regard the price rise as temporary and expect prices to fall later on. Consequently, they tend to increase their money holdings and the price rise is less than the rise in nominal money supply. Gradually people tend to readjust their money holdings. Price then rise more than in proportion to the money supply. The precise rate at which prices rise for a given rate of increase in the money supply depends on such factors as past price behaviour, current changes in the structure of labour, product markets and fiscal

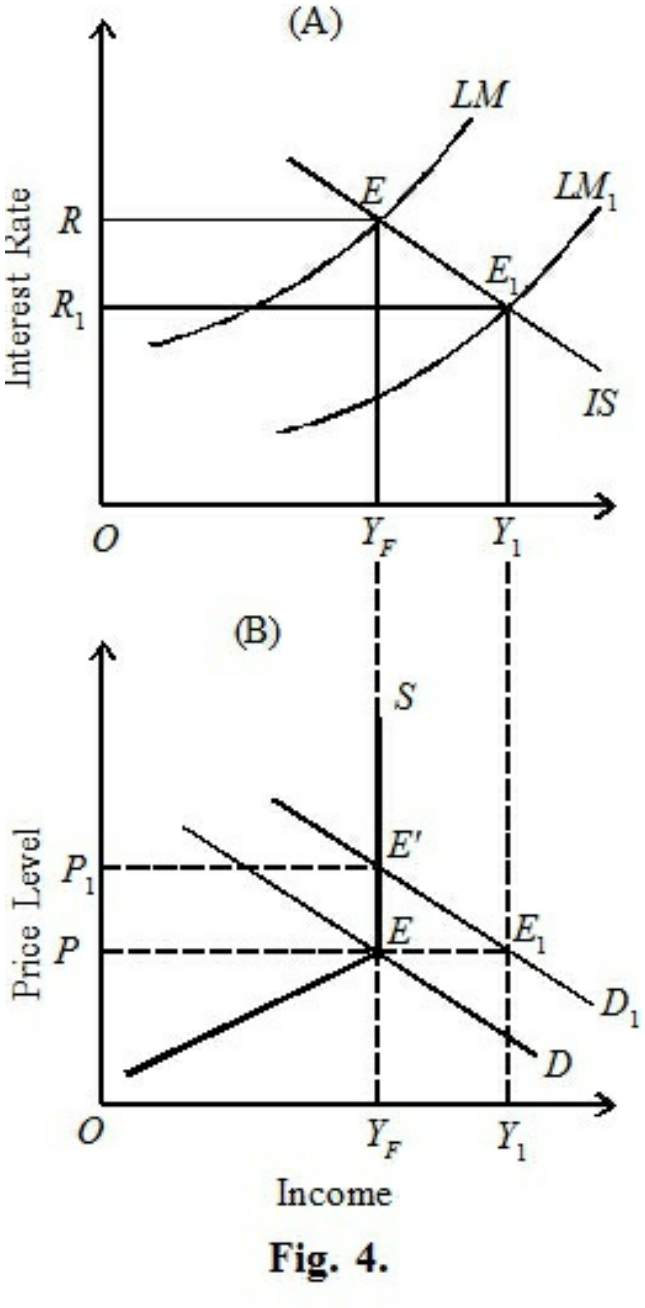


Fig. 4.

policy. Thus, according to Friedman, the monetary expansion works through output before inflation starts.

The quantity theory version of the demand-pull inflation is illustrated diagrammatically in Figure 4 (A) & (B). Suppose the money supply is increased at a given price level *P* as determined by *D* and *S* curves in Panel (B) of the figure. The initial full employment situation at this price level is shown by the intersection of IS and *LM* curves at *E* in Panel (A) of the figure where *R* is the interest rate and Y_F is the full employment level of income. Now with the increase in the quantity of money, the *LM* curve shifts rightward to LM_1 and intersects the *IS* curve at E_1 such that the equilibrium level of income rises to Y_1 and the rate of interest is lowered to R_1. As the aggregate supply is assumed fixed, there is no change in the position of the *IS* curve.

Consequently, the aggregate demand rises which shifts the *D* curve to the right to D_1 and thus excess demand is created equivalent to EE_1 ($=Y_FY_1$) in Panel (B) of the figure. This raises the price level, the aggregate supply being fixed, as shown by the vertical portion of the supply curve *S*. The rise in the price level reduces the real value of the money supply so that the LM_1 curve shifts to the left to *LM*. Excess demand will not be eliminated until aggregate demand curve D_1 cuts the aggregate supply curve *S* at *E'*. This means a higher price level P_1 in Panel (B) and return to the original equilibrium position *E* in the upper Panel of the figure where the *IS* curve cuts the *LM* curve. The "result, then is self-limiting, and the price level rises in exact proportion to the real value of the money supply to its original value."[14]

2. Keynes' Theory of Demand-Pull Inflation

Keynes and his followers emphasise the incease in aggregate demand as the source of demand-pull inflation. There may be more than one source of demand. Consumers want more goods and services for consumption purposes. Businessmen want more inputs for investment. Government demands more goods and services to meet civil and military requirements of the country. Thus the aggregate demand comprises consumption,

investment and government expenditures. When the value of aggregate demand exceeds the value of aggregate supply at the full employment level, the inflationary gap arises. The larger the gap between aggregate demand and aggregate supply, the more rapid the inflation. Given a constant average propensity to save, rising money incomes at the full employment level would lead to an excess of aggregate demand over aggregate supply and to a consequent inflationary gap. Thus Keynes used the notion of the inflationary gap to show an inflationary rise in prices.

The Keynesian theory is based on a short-run analysis in which prices are assumed to be fixed. In fact, prices are determined by non-monetary forces. On the other hand, output is assumed to be more variable which is determined largely by changes in investment spending. The Keynesian chain of causation between changes in nominal money income and in prices is an indirect one through the rate of interest. When the quantity of money increases, its first effect is on the rate of interest which tends to fall. A fall in the interest rate would, in turn, increase investment which would raise aggregate demand. A rise in aggregate demand would first affect only output and not prices so long as there are unemployed resoures. But a sudden large increase in the aggregate demand would encounter bottlenecks when resources are still unemployed.

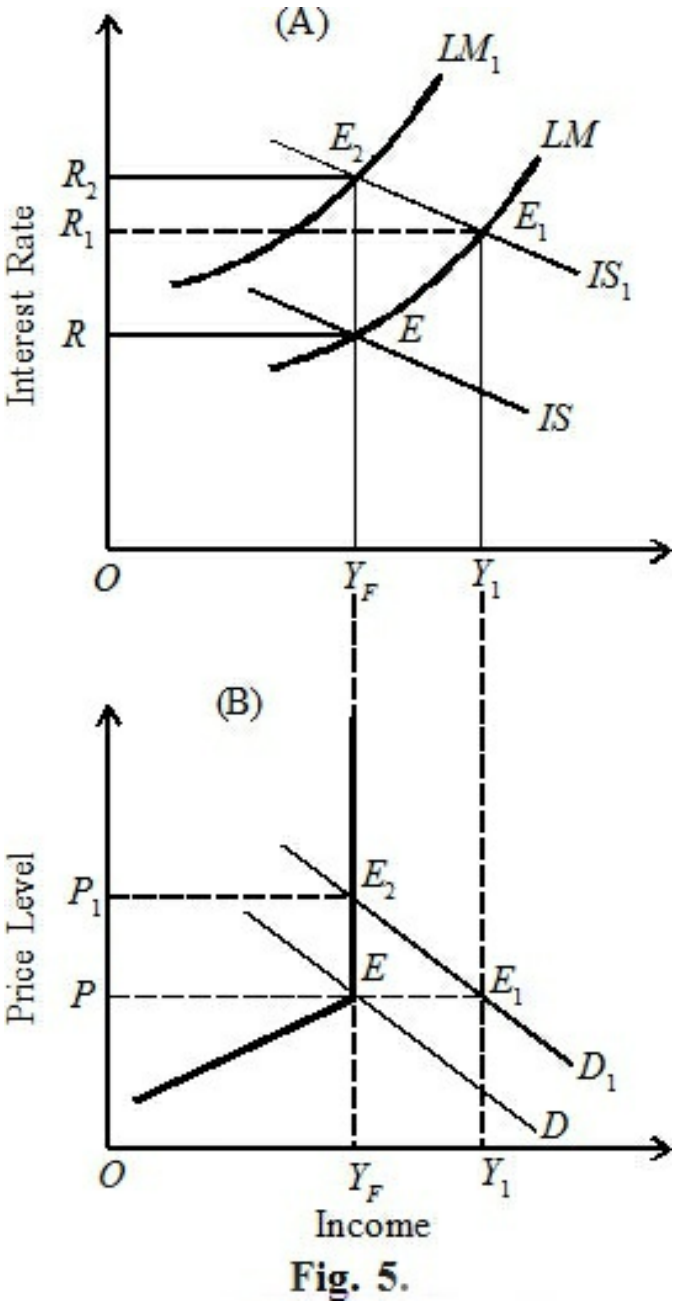


Fig. 5.

The supply of some factors might become inelastic or others might be in short supply and non-substitutable. This would lead to increase in marginal costs and hence in prices. Accordingly prices would rise above average unit cost and profits would increase rapidly which, in turn, would bid up wages owing to trade union pressures. Diminishing returns might also set in some industries. As full employment in reached, the elasticity of supply of output falls to zero and prices rise without any increase in output. Any further increase in expenditure would lead to excess demand and to more than proportional increase in prices. Thus, in the Keynesian view so long as there in unemployment, all the change in income is in output, and once there is full employment, all is in prices.

---

14. T. F. Dernberg and D.M. McDougall, op.cit., p. 291.

The Keynesian theory of demand-pull inflation is explained diagrammatically in Figure 5 (A) and (B). Suppose the economy is in equilibrium at *E* where the *IS* and *LM* curves intersect with full employment income level Y_F and interest rate *R*, as shown in Panel (A) of the figure. Corresponding to this situation, the price level is *P* in Panel (B). Now the government increases its expenditure. This shifts the *IS* curve rightward to IS_1 and intersects the *LM* curve at E_1 when the level of income and the interest rate rise to Y_1 and R_1 respectively. The increase in government expenditure implies an increase in aggregate demand which is shown by the upward shift of the *D* curve to D_1 in the lower Panel (B) of the figure. This creates excess demand to the extent of EE_1 ($=Y_FY_1$) at the initial price level *P*. Excess demand tends to raise the price level, as aggregate supply of output cannot be increased after the full employment level. As the price level rises, the real value of money supply falls. This shifts the *LM* curve to the left to LM_1 such that it cuts the IS_1 curve at E_2 where equilibrium is established at the full employment level of income Y_F, but at a higher interest rate R_2 (in Panel A) and a higher price level P_1 (in Panel B).

Thus the excess demand caused by the rise in government expenditure eliminates itself by changes in the real value of money.

3. Bent Hansen's Excess Demand Model

The Danish economist Bent Hansen[15] has presented an explicit dynamic excess demand model of inflation which incorporates two separate price levels, one for the goods market and other for the factor (labour) market.

Its Assumptions

His dynamic model for demand inflation is based on the following assumptions:

1. There is perfect competition in both the goods market and the factor market.

2. Price at the moment will persist in the future.

3. Only one commodity is produced with the help of only one variable factor, labour services.

4. The quantity of labour services per unit of time is a given magnitude.

5. There is a fixed actual level of employment and consequently of output which is full employment.

The Model

Given these assumptions, the model is explained in terms of Figure 6. The vertical axis measures the price-wage ratio P/W (inverse of the real wage). The aggregate real income or output is measured along the horizontal axis. S is the supply curve of planned production, $S=F\ (P/W)$. It varies positively with P/W such that the higher the price is relative to the wage rate, the less is the demand for consumer goods, $D=F(P/W)$. D is the demand curve of planned demand which has an inverse relationship with P/W such that the higher price is relative to the wage rate the larger is the planned production. The vertical line Q is the full employment output level Q_F and Q=constant.

The horizontal difference between the curve D and Q is the "quantitative

inflationary gap in the goods markets". Such a gap exists at all price-wage ratios below (P/W) in the figure. The horizontal difference between the curves S and Q is the index for the factor-gap." *Thus* ($D–Q$) *is the goods gap and* ($S–Q$) *is the factor gap.*

Suppose the two curves D and S intersect to the right of the full employment level of output at point E. This happens if there is monetary pressure of inflation because otherwise it would not be possible with given P/W to have a positive inflationary gap in the goods markets and positive factor-gap simultaneously. A monetary pressure of inflation exists only when P/W is between P/W_1 and P/W_4. When $P/W>P/W_1$, the inflationary gap in the goods-market is greater than zero; and when $P/W<P/W_4$ both the index for the factor-gap and the factor-gap are negative.

Next Hansen introduces two dynamical equations :

$$dp/dt = f\,(D–Q \qquad \text{...(1)}$$

$$dw/dt = F\,(S–Q \qquad \text{...(2)}$$

where dp/dt is the speed of the rise in the price level, and dw/dt is the speed of the rise in the wage rate.

When ($D–Q$) is zero, $dp/dt = 0$; and when ($S–Q$) is zero, dw/dt = O. This is a static equilibrium system. When the two gaps are positive, the rates of price and wage changes are also positive.

It follows that when both the excess demand for goods ($D–Q$) and the excess demand for factors ($S–Q$) are positive, both price and wage-rate will rise. Each will be a quasi-equilibrium position which is stable in the sense that whatever price-wage relation is started, there will be forces at work which tend to bring the system back to the quasi-equilibrium position.

The quasi-equilibrium system is given by

$$Q = \text{Constant} \qquad S = F(P/W) \qquad D = f(P/W)$$

and $$P/W = \frac{f(D-Q)}{F(S-Q)}$$

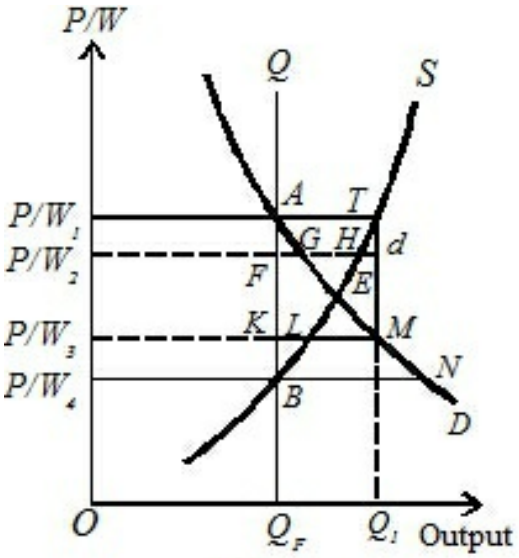


Fig. 6.

Let us take the figure where the curves *S* and *D* intersect at point *E*, to the right of the full employment level of output Q_F. Since point *E* cannot be achieved, an initial unstable equilibrium occurs at point *A* where the price-wage ratio is (P/W_1). In this situation, there is no goods gap and goods prices do not rise because planned demand (*D*) equals full employment output (Q_F) at *A*. But there is a large factor gap at point *T* so that wages rise rapidly. This is because planned production Q_1 exceeds the full employment output Q_F at (P/W_1). But this is not possible because Q_1 output is more than the full employment output Q_F. Consequently, there is excess demand for labour which leads to labour shortages and to rise in the wage rate. Thus *P*/*W* falls. When the price-wage ratio falls, an excess demand for goods (goods gap) begins to appear and that for factors (factor gap) simultaneously decreases. Suppose P/W_1 falls to P/W_2. At P/W_2, the goods gap *FG* is smaller than the factor gap *FH* which means that the small goods gap produces a slow rise in prices and the larger factor gap produces a higher rise in wage rate. This will lead to a further fall in wage-price ratio to P/W_3. At P/W_3, the factor gap is reduced to *KL* and the goods gap is raised to *KM*, thereby leading to a slower rise in wage rate and more rapid rise in prices respectively. This retards the fall in the wage-price ratio. In this way, the price-wage ratio will fall, increasing slowly to a level where the goods gap corresponds to the factor gap. This means that the precentage rise of wage rate per unit time is equal to the percentage rise of the price per unit time. Similar reasoning will apply if we start from P/W_4 where the large goods gap *BN* and zero factor gap would raise prices and hence the wage-price ratio. A key determinant of the level of price-wage ratio is the

flexibility of wage-rate and prices relative to each other. The more flexible are prices relative to wages, the closer is the value of price-wage ratio to P/W_1.

15. Bent Hansen, *A study in the theory of Inflation*, 1951, Ch. 7.

Between P/W_1 and P/W_4, there is some quasi-equilibrium at which both prices and wage-rates move together. The quasi-equilibrium is not a static equilibrium but a dynamic one, since both prices and wage rates rise without interruption and the relevant gaps are not zero. "The actual speed of the inflation to quasi-equilibrium will depend on the absolute sensitivity of wage and price change to the size of the relevant gaps. If both are relatively volatile, inflation will be rapid; if both are relatively sluggish, inflation will be slower." The more rigid prices are relative to wages, the closer is the value of price-wage ratio to P/W_4.

To conclude, Hensen's excess demand model of inflation points toward the sources of inflationary pressures and the actual process of inflation in the economy. But, according to Ackley, it fails to specify the rate at which inflation will occur. It is an elegant but perhaps rather empty analysis of demand inflation.

COST-PUSH INFLATION

Cost-push inflation is caused by wage increases enforced by unions and profit increases by employers. The type of inflation has not been a new phenomenon and was found even during the medieval period.[16] But it was revived in the 1950s and again in the 1970s as the principal cause of inflation. It also came to be known as the "*New Inflation*."[17] Cost-push inflation is caused by wage-push and profit-push to prices.

The basic cause of cost-push inflation is the rise in money wages more rapidly than the productivity of labour. In advanced countries, trade unions are very powerful. They press employers to grant wage increases considerably in excess of increases in the productivity of labour, thereby raising the cost of production of commodities. Employers, in turn, raise

prices of their products. Higher wages enable workers to buy as much as before, in spite of higher prices. On the other hand, the increase in prices induces unions to demand still higher wages. In this way, the wage-cost spiral continues, thereby leading to cost-push or wage-push inflation.

Cost-push inflation may be further aggravated by upward adjustment of wages to compensate for rise in the cost of living index. This is usually done in either of the two ways. *First*, unions include an "escalator clause" in contracts with employers, whereby money wage rates are adjusted upward each time the cost of living index increases by some specified number of percentage points. *Second*, in case where union contracts do not have an escalator clause, the cost of living index is used as the basis for negotiating larger wage increases at the time of fresh contract settlements.

Again, a few sectors of the economy may be affected by money wage increases and prices of their products may be rising. In many cases, their products are used as inputs for the production of commodities in other sectors. As a result, production costs of other sectors will rise and thereby push up the prices of their products. Thus wage-push inflation in a few sectors of the economy may soon lead to inflationary rise in prices in the entire economy.

Further, an increase in the price of domestically produced or imported raw materials may lead to cost-push inflation. Since raw materials are used as inputs by the manufacturers of finished goods, they enter into the cost of production of the latter. Thus a continuous rise in the prices of raw materials tends to set off a cost-price-wage spiral.

Another cause of cost-push inflation is *proft-push inflation*. Oligopolist and monopolist firms raise the prices of their products to offset the rise in labour and production costs so as to earn higher profits. There being imperfect competition in the case of such firms, they are able to "administer prices" of their products. "In an economy in which so called administered prices abound, there is at least the possibility that these prices may be administered upward faster than costs in an attempt to earn greater profits. To the extent such a process is widespread profit-push inflation will result."[18] Profit-push inflation is, therefore, also called *administered-price theory of inflation* or *price-push inflation* or *sellers*'

inflation or *market-power inflation.*

But there are certain limitations on the power of firms to raise their profits. They cannot raise their selling prices to increase their profit-margins if the demand for their products is stable. Moreover, firms are reluctant to increase their profits every time unions are successful in raising wages. This is beause profits of a firm depend not only on price but on sales and unit costs as well, and the latter depend in part on prices charged. So firms cannot raise their profits becuase their motives are different from unions. *Lastly*, profits form only a small fraction of the price of the product and a once-for-all increase in profits is not likely to have much impact on prices. Economists, therefore, do not give much importance to profit-push inflation as an explanation of cost-push inflation.

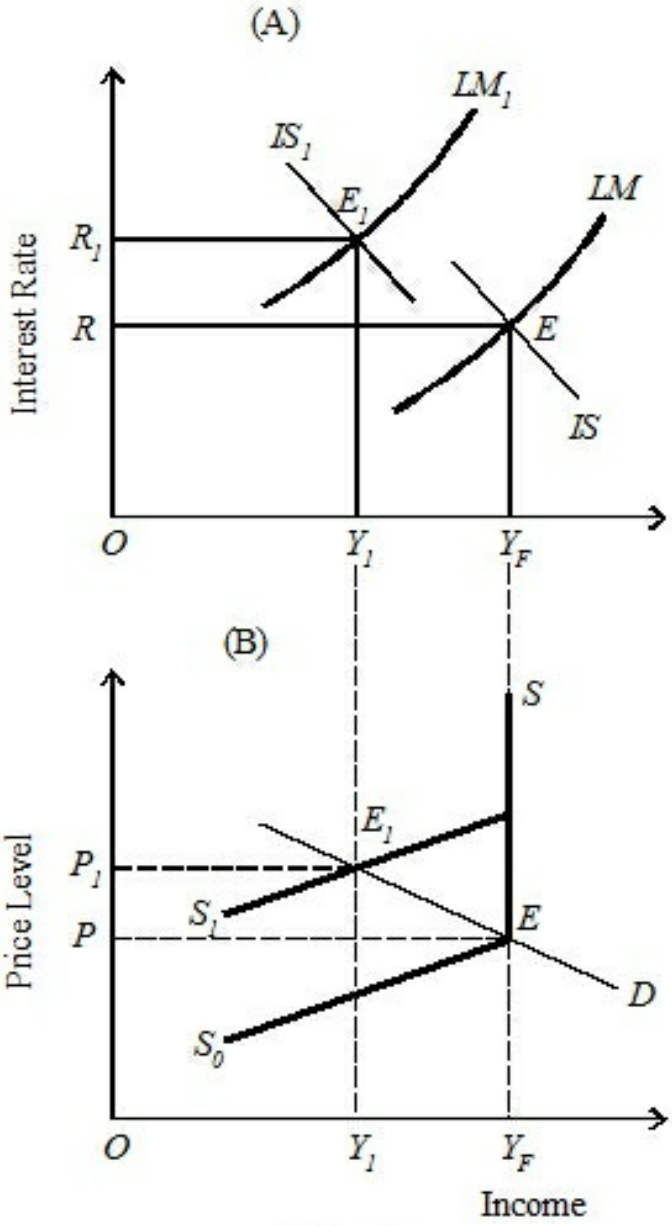


Fig. 7.

Cost-push inflation is illustrated in Figure 7 (A) and (B). First consider Panel (B) of the figure where the supply curves S_OS and S_1S are shown as increasing functions of the price level up to full employment level of income Y_F. Given the demand conditions as represented by the demand curve D, the supply curve S_O is shown to shift to S_1 in response to cost-increasing pressures of oligopolies, unions, etc. as a result of rise in

money wages. Consequently, the equilibrium position shifts from E to E_1 reflecting rise in the price level from P to P_1 and fall in output, employment and income from Y_F to Y_1 level.

---

16. For a brief survey of post-War developments in inflation theories see M. Bronfenbrenner and F.D. Holzman, "Survey of Inflation Theory," *AER*, Sept. 1963.

17. It is the title of a book by W.L. Thorpe and R.E. Quandt, 1959.

Now consider the upper Panel (A) of the figure. As the price level rises, the *LM* curve shifts to the left to LM_1 because with the increase in the price level to P_1 the real value of the money supply falls. Similarly, the *IS* curve shifts to the left to IS_1 because with the increase in the price level, the demand for consumer goods falls due to the Pigou effect. Accordingly, the equilibrium position of the economy shifts from E to E_1 where the interest rate increases from R to R_1 and the output, employment and income levels fall from the full employment level of Y_F to Y_1.

Its Criticisms

The cost-push theory has been criticised on three issues.

First, cost-push inflation is associated with unemployment. So the monetary authority is in a fix because to control inflation it will have to tolerate unemployment.

Second, if the government is committed to a policy of full employment, it will have to tolerate wage increases by unions, and hence inflation.

Lastly, if the government tries to increase aggregate demand during periods of unemployment, it may lead to increase in wages by trade union action instead of raising output and employment.

DEMAND-PULL VERSUS COST-PUSH INFLATION

There has been a lot of controversy among economists over the issue

whether inflation is the consequence of demand-pull or cost-push. According to F. Machlup, "The distinction between cost-push and demand-pull inflation is unworkable, irrelevant or even meaningless."[19]

However, the debate between demand-pull and cost-push inflation arises mainly from the difference between the policy recommendations on the two views. Recommendations on demand-pull inflation are related to monetary and fiscal measures which lead to a higher level of unemployment. On the other hand, recommendations on cost-push aim at controlling inflation without unemployment through administrative controls on price increases and incomes policy.

---

18. Edward Shapiro, *op. cit.,* p. 522.

Machlup argues that the controversial issue is partly who is to be blamed for inflation and partly what policies should be pursued to aviod a persistent increase in prices. If demand-pull is the cause of inflation then the government is blamed for overspending and taxing little, and the central bank is blamed for keeping interest rates too low and for expansion of too much credit. On the other hand, if cost-push is the cause of inflation then trade unions are blamed for excessive wage increases, industry is blamed for granting them, big firms for raising administered prices of materials and goods to earn higher profits and government is blamed for not persuading or forcing unions and industry from raising their wages and profits. But trade unions reject the wage-push theory because they would not like to be blamed for inflation. They also reject the demand-pull view because that would prevent the use of monetary and fiscal measures to increase employment. Thus they hold only big firms responsible for inflationary rise in prices through administered prices. But there is no conclusive proof that the profit margins and profit rates of firms have been increasing year after year.

Machlup further points out that there is a group of economists who holds that cost-push is no cause of inflation, "because, without an increase in purchasing power and demand, cost increases would lead to unemployment." On the other hand, there is another group of economists who believes that demand-pull is no cause of inflation, it takes a cost-push

to produce it.

Thus it is difficult to distinguish demand-pull from cost-push inflation in practice and it is easy to say that inflation has been caused by cost-push when, in fact, demand-pull may be the cause. As pointed out by Samuelson and Solow, "The trouble is that we have no normal initial standard from which to measure, no price level which has always existed to which every one has adjusted."[20] It is also suggested that identification of demand-pull or cost-push inflation can be made with reference to timing. If prices increase first, it is a demand-pull inflation, and if wages increase follow, it is a cost-push inflation.

Like Machlup, Johnson regards the issue of demand-pull versus cost-push as "largely a spurious one." He assigns three reasons for this, *First,* the proponents of the two theories fail to investigate the monetary assumptions on which the theories are based. Neither the demand-pull nor the cost-push theory can generate a sustained inflation unless monetary policy followed by the monetary theory is taken into consideration under varying circumstances. The two theories are, therefore, not independent and self-contained. The *second* reason is based on differences between the two theories about their definitions of full employment. If full employment is defined as a situation when the demand for goods is just sufficient to prevent from rising or falling, then it is a case of demand-pull inflation which is associated with excess demand for goods and labour. Full employment here means overfull employment. On the other hand, if full employment is defined as the level of unemployment at which the percentage of the unemployment just equals the number of persons seeking jobs, then inflation is caused by forces other than excess demand. Such forces cause cost-push inflation. In the *third* place, it is extremely difficult to devise a test capable of determining whether a particular inflation is of the demand-pull or cost-push type.[21]

We may conclude with Lipsey :"Debate continues on the balance between demand and cost as forces causing inflation in the contermporary inflationary climate. The debate is important because the policy implications of different causes of inflation are different, and different target variables need to be controlled, according to the cause. Until the

causes of inflation are fully understood, there will be debate about policies."[22]

---

19. F. Machlup, "Another View of Cost-Push and Demand-Pull Inflation," *R.E.S.* Vol. 42, May 1960

MIXED DEMAND-PULL COST-PUSH INFLATION

Some economists do not accept this dichotomy that inflation is either demand-pull or cost-push. They hold that the actual inflationary process contains some elements of both. In fact, excess demand and cost-push forces operate simultaneously and interdependently in an inflationary process. Thus inflation is mixed demand-pull and cost-push when price level changes reflect upward shifts in both aggregate demand and supply functions.

But it does not mean that both demand-pull and cost-push inflations may start simultaneously. In fact, an inflationary process may begin with either excess demand or wage-push. The timing in each case may be different. In demand-pull inflation, price increases may precede wage increases, while it may be the other way round in the case of cost-push inflation. So price increases may start with either of the two forces, but the inflationary process cannot be sustained in the absence of the other forces.

Suppose an inflationary process begins with excess demand with no cost-push forces at work. Excess demand will raise prices which will in due course pull up money wages. But the rise in money wages is not the result of cost-push forces. Such a mixed inflation will lead to sustained rise in prices. This is illustrated in Figure 8. The initial equilibrium is at Y_F level of full employment income determined by aggregate demand D_0 and aggregate supply S_oS curves at A. The price level is P_o with increase in aggregate demand from D_o to D_1 and D_2 given the vertical portion of the supply curve S_oS, prices rise from P_o to P_2 to P_5 , the inflationary path being A, B and C. This sustained increase in prices has also been the result of the increase in money wage rates due to increase in aggregate demand at the full employment level. When prices rise, producers are

encouraged to increase output as their profits rise with increased aggregate demand. They, therefore, raise the demand for labour thereby increasing money wages which further lead to increase in demand for goods and services. So long as the demand for output continues to raise money incomes, inflationary pressures will continue.

Consider an inflationary process that may begin from the supply side due to increase in money wage rates. This will raise prices every time there is a wage-push. But the rise in prices will not be sustained if there is no increase in demand. This is illustrated in Fig. 8 where given the aggregate demand curve D_o, a wage-push shifts the supply curve S_o to S_1. The new equilibrium is at E. This raises the price level from P_o to P_1 and lowers output and employment to Y_2 below the full employment level Y_F. A further wage-push will again shift the supply curve to S_2, and the new equilibrium will be at F, given the demand curve D_o thereby raising the price level further to P_3 and also reducing output and employment to Y_1. In the absence of increase in aggregate demand, this cost-push inflationary process will not be a sustained one and will sooner or later come to an end.

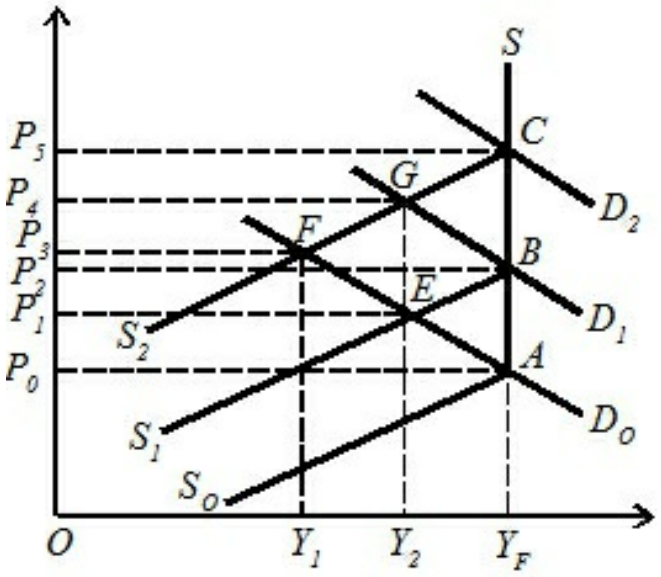


Fig. 8

---

20. "Analytical Aspects of Anti-Inflationary Policy." *A.E.R.*, Vol. 50, 1960.

21. H. G. Johnson, *Op, cit*, pp. 128-30.

22. Richard G. Lipsey, *Introduction to Positive Economics*, 8/e, 1995

The cost-push inflationary process will be self-sustaining only if every wage-push is accompanied by a corresponding increase in aggregate demand. Since every cost-push is accompained by a fall in output and employment along with a price increase, it is likely that the government

will adopt expansionary monetary and fiscal policies in order to check the fall in output and employment. In this way, cost-push will lead to a sustained inflationary process because the government will try to achieve full employment by raising aggregate demand which will, in turn, lead to further wage-push and so on. Such a situation is again explained with the help of Figure 8. Suppose there is a wage-push at E which shifts the supply curve from S_1 to S_2 and equilibrium is established at F with the demand curve D_o. The price level rises to P_3 and the level to employment is reduced to Y_1. When due to an expansionary monetary and fiscal policy, aggregate demand increases to D_1, the new equilibrium position is at G where the price level rises to P_4 and the level of employment rises to Y_2. A further increase in demand shifts the aggregate demand curve upward to D_2 , such that equilibrium is attained at point C where the price level rises to P_5 and the economy attains the full employment level Y_F. Thus a wage-push accompanied by an increase in aggregate demand through expansionary monetary and fiscal policies traces out a ratchet-like inflationary parth from A to E to F to G and to C.

SECTORAL OR DEMAND-SHIFT INFLATION

Sectoral or demand-shift inflation is associated with the name of Charles Schulz[23] who in a paper, pointed out that the price increases from 1955-57 were caused by neither demand-pull nor cost-push but by sectoral shifts in demand. Schultz advanced his thesis with reference to the American economy but it has now been generalised in the case of modern industrial economies.

Schultz begins his theory by pointing out that prices and wages are flexible upward in response to excess demand but they are rigid downward. Even if the aggregate demand is not excessive, excess demand in some sectors of the economy and deficient demand in other sectors will still lead to a rise in the general price level. This is because prices do not fall in the deficient-demand sectors, there being downward rigidity of prices. But prices rise in the excess-demand sectors and remain constant in the other sectors. The net effect is an overall rise in the price level.

Moreover, increase in prices in excess-demand industries (or sectors) can spread to deficient-demand industries through the prices of materials and the wages of labour. Excess demand in particular industries will lead to a general rise in the price of intermediate materials, supplies and components. These rising prices of materials will spread to demand-deficient industries which use them as inputs. They will, therefore, raise the prices of their products in order to protect their profit margins.

Not only this, wages will also be bid up in excess demand industries, and wages in demand-deficient industries will follow the rising trend. Because if wages in the latter industries are not raised, they will lead to dissatisfaction among workers, thereby leading to inefficiency and fall in productivity. Thus rising wage rates, originating in the excess demand industries, spread throughout the economy.

The spread of wage increases from excess demand industries to other parts of the economy increases the rise in the price of semi-manufactured materials and components. Other things remaining the same, the influence of increasing costs will be larger at the final stages of production. Thus producers of finished goods will face a general rise in the level of costs, thereby leading to rising prices. This may happen even in case of those industries which do not have excess demand for the products.

Another reason for demand-shift inflation in modern industrial economies is increase in the relative importance of *overhead costs*. This increase is due to two factors. *First*, there is an increase in overhead staff at the expense of production workers. According to Schultz, automation of production methods, instrumentation of control functions, mechanisation of office and accounting procedures, self-regulating materials, handling equipments, etc. lead to the growth of professional and semi-professional personnel in supervising, operating and maintenance roles. Similarly, the growth of formal research and development (R&D) as a separate function not only alters the production processes but also the composition of the labour force required to service them. These developments lead to the decline in the ratio of production workers to technical and supervisory staff in industries.

The *second* reason for the rise in overhead costs is that the ratio of

relatively short-lived equipment to long-lived plant rises substantially. As a result, depreciation as a proportion of total cost increases. The ultimate effect of an increasing proportion of overhead costs in the total cost is to make average costs more sensitive to variations in output. The distinguishing characteristic of the demand-shift inflation is a continued investment boom in the face of stable aggregate output. All industries expand their capacity and their employment of overhead personnel yet only a few enjoy a concomitant rise in sales. So producers facing shrinking profit margins try to recover a part of their rising costs in higher prices.

Thus demand-shift inflationary process "arises initially out of excess demand in particular industries. But it results in a general price rise only because of the downward rigidities and cost-oriented nature of prices and wages. It is not characterized by an autonomous upward push of costs nor by an aggregate excess demand. Indeed its basic nature is that it cannot be understood in terms of aggregates alone. Such inflation is the necessary result of sharp changes in the composition of demand, given the structure of prices and wages in the economy."

This theory was evolved by Schultz to examine the nature of the gradual inflation to which the American economy had been subjected during the period 1955-57. It has since been generalised in the case of modren industrial economies.

Its Criticisms

Johnson has criticised this theory for two reasons.

First, empirical evidence has failed to confirm Schultz's proposition that sectoral price increases are explained by upward shifts of demand.

Second, it suffers from the same defects as the two rival theories of demand-pull and cost-push, it seeks to challenge. That is, its "failure to investigate the monetary preconditions for inflation, and imprecision respecting the definitions of full employment and general excess demand."

23. C.L. Shultz, "Recent Inflation in the United States", in *Employment, Growth and Price Level*, Study Paper 1, Joint Economic Committee, U.S. Congress, 1959.

STRUCTURAL INFLATION

The structuralist school of South America stresses structural rigidities as the principal cause of inflation in such developing countries as Argentina, Brazil, and Chile, Of course, this type of inflation is also to be found in other developing countries.

The structuralists hold the view that inflation is necessary with growth. According to this view, as the economy develops, rigidities arise which lead to structural inflation. In the initial phase, there are increases in non-agricultural incomes accompanied by high growth rate of population that tend to increase the demand for goods. In fact, the pressure of population growth and rising urban incomes would tend to raise through a chain reaction mechanism, *first* the prices of agricultural goods, *second*, the general price level, and *third*, wages. Let us analyse them.

1. Agricultural Goods. As the demand for agricultural goods rises, their domestic supply being inelastic, the prices of agricultural goods rise. The output of these goods does not increase when their prices rise because their production is inelasitc due to a defective system of land tenure and other rigidities in the form of lack of irrigation, finance, storage and marketing facilities, and bad harvests. To prevent the continuous rise in agricultural products, especially foods products, they can be imported. But it is not possible to import them in large quantities due to foreign exchange constraint. Moreover, the prices of imported products are relatively higher than their domestic prices. This tends to raise the price level further within the economy.

2. Wage Increases. When the prices of food products rise, wage earners press for increase in wage rates to compensate for the fall in their real incomes. But wages and/or D.A. are linked to the cost of living index. They are, therefore, raised whenever the cost of living index rises above an agreed point which further increases the demand for goods and a further rise in their prices.

The effect of increase in the wage rates on prices in illustrated in Figure 9. When wage rates rise, the aggregate demand for goods increases from D_1 to D_2. But the aggregate supply falls due to increase in labour costs which results in the shifting of aggregate supply curve from S_1S to S_2S. Since the production of goods is inelastic due to structural rigidities after a point, the supply curve is shown as vertical from point E_1 onward. The initial equilibrium is at E_1 where the curves D_1 and S_1 intersect at the output level OY_1 and the price level is OP_1. When supply falls due to increase in labour costs, the supply curve shifts from S_1 to S_2 and it intersects the demand curve D_2 at E_2 and production falls from OY_1 to OY_2 and the price level rises from OP_1 to OP_2.

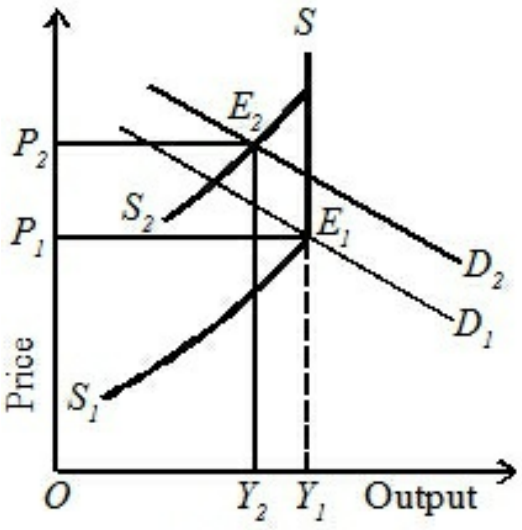


Fig. 9

3. Import Substitution. Another cause of structural inflation is that the rate of export growth in a developing economy is slow and unstable which is inadequate to support the required growth rate of the economy. The sluggish growth rate of exports and the foreign exchange constraint lead to the adoption of the policy of industrialisation based on import substitution. Such a policy necessitates the use of protective measures which, in turn, tend to raise the prices of industrial products, and incomes in the non-agricultrural sectors, thereby leading to further rise in prices. Moreover, this policy leads to a cost-push rise in prices because of the rise in prices of imported materials and equipment, and protective measures. The policy of import substitution also tends to be inflationary because of the relative inefficiency of the new industries during the "learning" period. The secular deterioration in the terms of trade of primary products of developing countries further limits the growth of income from exports which often leads to the exchange rate devaluation.

4. Tax System. The nature of the tax systems and budgetary processes also help in accentuating the inflationary trends in such economies. The

tax system has low inflation elasticities which means that when prices rise, the real vlaue of taxes falls. Often taxes are fixed in money terms or they are raised slowly to adjust for price rises. Moreover, it often takes long time to collect taxes with the result that by the time they are paid by assessees, their real value is less to the exchequer. On the other hand, planned expenditures on projects are often not incurred on schedule due to various supply bottlenecks with the result that when prices rise, the money value of expenditures rises proportionately. As a result of fall in the real value of tax collections and rise in money value of expenditures, governments have to adopt larger fiscal deficits which further accentuate inflationary pressures.

5. Money Supply. So far as the money supply is concerned, it automatically expands when prices rise in a developing country. As prices rise, firms need larger funds from banks. And the government needs more money to finance larger deficits in order to meet its expanding expenditure and wages of its employees. For this, it borrows from the central bank which leads to monetary expansion and to a further rise in the rate of inflation.

Thus structural inflation may result from supply inelasticities leading to rise in agricultural prices, costs of import substitutes, deterioration of the terms of trade and exchange rate devaluation.

Its Criticisms

The basic weaknesses in the structural arguments have been :

First, no separation is made between autonomous structural rigidities and induced rigidities resulting from price and exchange controls or mismanagement of government intervention.

Second, the sluggishness in the export growth is not really structural but the result of failure to exploit export opportunities because of overvalued exchange rates.

MARKUP INFLATION

The theory of markup inflation is mainly associated with Prof. Ackley, though formal models have also been presented by Holzman and Duesenberry independently of each other.[24] We analyse below Ackley's simplified version of the markup inflation.

The analysis is based on the assumption that both wages and prices are "administered" and are settled by workers and business firms. Firms fix administrative prices for their goods by adding to their direct material and labour costs, and some standard markup which covers profit.* Labour also seeks wages on the basis of a fixed markup over its cost of living.

This model of inflation can lead to either a stable, a rising, or a falling price level depending on the markups which firms and workers respectively use. If either or both use a percentage markup, the inflation will progress faster than if either or both fix the markups in money terms. If each participant fixes prices on the basis of prices he pays, the inflation will be high and of long duration. If one firm raises its prices in order to maintain its desired markup, the costs of other firms are raised which, in turn, raise their prices and this process of raising costs and prices will spread to other firms in an endless chain. When consumers buy such goods whose prices are rising, their cost of living rises. This causes wage costs to rise, thereby increasing the inflationary spiral. However, the inflationary spiral may come to a halt, if there is a gradual improvement in the efficiency and productivity of labour. A rise in efficiency and productivity means that there is a rise in wage rates or prices of materials leading to a smaller rise in labour and material costs. But stability in prices may not come if firms and workers appropriate the gains of rising productivity by increasing their markups. If each participant increases its markup by 100 per cent of the gains of productivity increase, the inflationary spiral might continue indefinitely.

According to Ackley, the markup can be based on either historical experience or expectations of future costs and prices. Moreover, the size of the markup applied by firms and workers is a function of the pressure of demand felt in the economy. When the demand is moderate, the markups may by applied to historically experienced costs and prices, and the price rise may be slow. But when demand is intense, the markups are

based on anticipations of future costs, and prices rise rapidly. Thus there can be no inflation without some change in the size of the markup.

This theory can also be applied to cost-push and demand-pull models of inflation. If firms and workers believe that their markups are lower than the required costs and prices, regardless of the state of aggregate demand, they will increase the size of their markups. Under such a situation, costs and prices rise in an inflationary spiral. This is similar to the cost-push inflation. On the other hand, if firms and workers raise the markups due to increase in demand, markup pricing is related to demand-pull inflation.

To conclude with Ackley, "Inflation might start from an initial autonomous increase either in business and labour markups. Or it might start from an increase in aggregate demand and which first and most directly affected some of the flexible market-determined prices. But however it starts, the process involves the interaction of demand and market elements."

The markup inflation can be controlled by the usual monetary and fiscal tools in order to restrict the demand for goods and increase productivity. Ackley also suggests wage-and-price guidelines or an incomes policy to be administered by a national wage-and-price commission.

Its Criticisms

Ackley's theory suffers from two weaknesses:

First, the theory gives a very limited explanation of the cause of inflation, especially the motives which compel workers and firms to fix higher markups in the absence of demand conditions.

Second, it suffers from the implication that once inflation starts, it is likely to continue indefinitely when costs and prices rise in a spiral.

Open and Suppressed Inflation

Inflation is often open and suppressed.

Inflation is open when “markets for goods or factors of production are allowed to function freely, setting prices of goods and factors without normal interference by the authorities.” Thus open inflation is the result of the uninterrupted operation of the market mechanism. There are no checks or controls on the distribution of commodities by the government. Increase in demand and shortage of supplies persists which tend to lead to open inflation. Unchecked open inflation ultimately leads to hyper-inflation.

---

24. G. Ackley, "A Third Approach to the Analysis and Control of Inflation", in *The Relationship of Prices to Economic Stability and Growth*, Joint Economic Committee, 1958; F.D. Holzman, "Income Determination in Open Inflation," and J. Duesenberry, "The Mechanics of Inflation," *R.E.S.*, May 1959.

* Markup covers both overhead costs and profit.

Suppressed Inflation. On the contrary when the government imposes physical and monetary controls to check open inflation, it is known as repressed or suppressed inflation. The market mechanism is not allowed to function normally by the use of licensing, price controls and rationing in order to suppress extensive rise in prices. According to Friedman, governments themselves are often producers and sellers of wide range of commodities and they want to keep their own prices low by price restrictions and controls. This leads to the breakdown of the free price system.

Further, suppressed inflation also results when efforts are made to increase domestic production and reduce import demand by tariffs, import restrictions, limits on foreign loans, voluntary import agreements, etc. So long as such controls exist, the present demand is postponed and there is diversion of demand from controlled to uncontrolled commodities. But as soon as these controls are removed, there is open inflation.

Its Effects

Suppressed inflation adversely affects the economy.

(1) When the distribution of commodities is controlled, the prices of uncontrolled commodities rise very high.

(2) Suppressed inflation reduces the incentive to work because people do not get the commodities which they want to have.

(3) Controlled distribution of goods also leads to malallocation of resources. This results in the diversion of productive resources from essential to non-essential industries.

(4) Frictions increase in the labour market when high inflation is associated with higher unemployment.

(5) Suppressed inflation leads to black marketing, corruption, hoarding and profiteering. It invites extra-legal powers of control.

(6) It reduces the prospect of anti-inflationary policy being tried at all.

THE PHILLIPS CURVE : THE RELATION BETWEEN UNEMPLOYMENT AND INFLATION

The Phillips curve examines the relationship between the rate of unemployment and the rate of money wage changes. Known after the British economist A.W. Phillips[25] who first identified it, it expresses an inverse relationship between the rate of unemployment and the rate of increase in money wages. Basing his analysis on data for the United Kingdom, Phillips derived the empirical relationship that when unemployment is high, the rate of increase in money wage rates is low. This is because "workers are reluctant to offer their services at less than the prevailing rates when the demand for labour is low and unemployment is high so that wage rates fall very slowly." On the other hand, when unemployment is low, the rate of increase in money wage rates is high. This is because, "when the demand for labour is high and there are very few unemployed we should expect employers to bid wage rates up quite rapidly."

The second factor which influences this inverse relationship between money wage rate and unemployment is the nature of business activity. In a period of rising business activity when unemployment falls with increasing demand for labour, the employers will bid up wages.

Conversely in a period of falling business activity when demand for labour is decreasing and unemployment is rising, employers will be reluctant to grant wage increases. Rather, they will reduce wages. But workers and unions will be reluctant to accept wage cuts during such periods. Consequently, employers are forced to dismiss workers, thereby leading to high rate of unemployment. Thus when the labour market is depressed, a small reduction in wages would lead to large increase in unemployment. Phillips concluded on the basis of the above arguments that the relation between rates of unemployment and a change of money wages would be highly non-linear when shown on a diagram. Such a curve is called the Phillips curve.

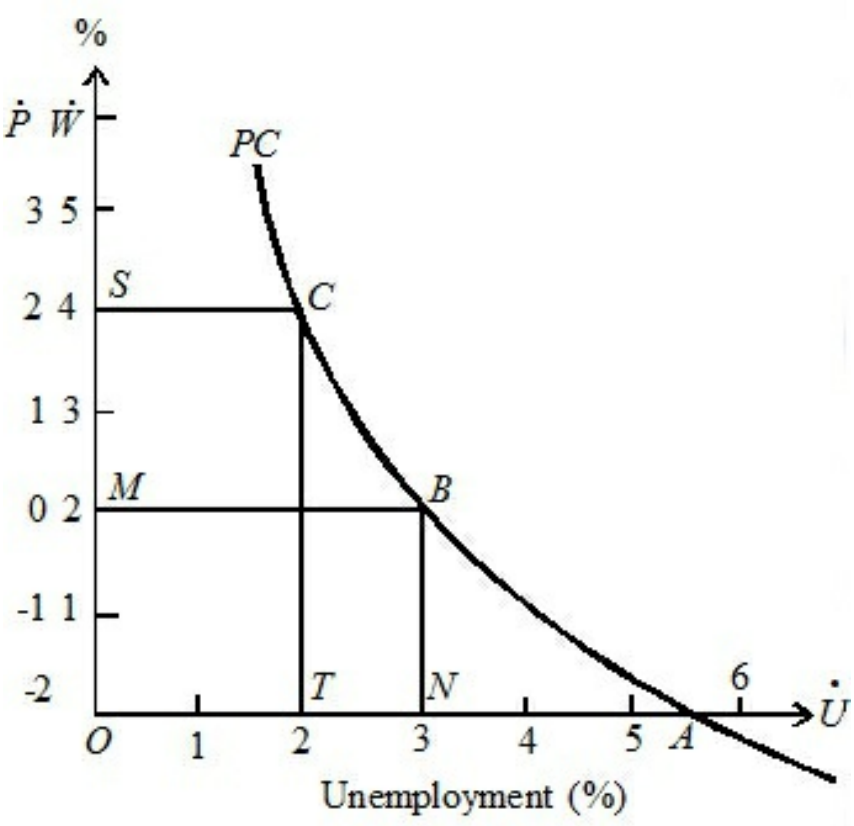


Fig. 10

The *PC* curve in Figure 10 is the Phillips curve which relates precentage change in money wage rate (*W*) on the vertical axis with the rate of unemployment (*U*) on the horizontal axis. The curve is convex to the origin which shows that the percentage change in money wages rises with decrease in the employment rate. In the figure, when the money wage rate is 2 per cent, the unemployment rate is 3 per cent. But when the wage rate is high at 4 per cent, the unemployment rate is low at 2 per cent. Thus there is a trade-off between the rate of change in money wage and the rate of unemployment. This means that when the wage rate is high the unemployment rate is low and vice versa.

The original Phillips curve was an observed statistical relation which was explained theoretically by Lipsey as resulting from the behaviour of labour market in disequilibrium through excess demand.[26]

Several economists have extended the Phillips curve analysis to the trade-off between the rate of unemployment and the rate of change in the level of prices or inflation rate by assuming that prices would change whenever wages rose more rapidly than labour productivity. If the rate of increase in money wage rates is higher than the growth rate of labour productivity, prices will rise and vice versa. But prices do not rise if labour productivity increases at the same rate as money wage rates rise.

This trade-off between the inflation rate and unemployment rate is explained in Figure 10 where the inflation rate $(\dot{P})$ is taken alongwith the rate of change in money wages $(\dot{W})$. Suppose labour productivity rises by 2 per cent per year and if money wages also increase by 2 per cent, the price level would remain constant. Thus point B on the *PC* curve corresponding to percentage change in money wages (*M*) and unemployment rate of 3 per cent (*N*) equals zero (*O*) per cent inflation rate $(\dot{P})$ on the vertical axis.* Now assume that the economy is operating at point *B*. If now, aggregate demand is increased, this lowers the unemployment rate to *OT* (2%) and raises the wage rate to *OS* (4%) per year. If labour productivity continues to grow at 2 per cent per annum, the price level will also rise at the rate of 2 per cent per annum at *OS* in the figure. The economy operates at point C. With the movement of the economy from *B* to *C*, unemployment falls to *T* (2%). If points *B* and *C* are connected, they trace out a Phillips curve *PC*.

Thus a money wage rate increase which is in excess of labour productivity leads to inflation. To keep wage increase to the level of labour productivity (*OM*) in order to avoid inflation. *ON* rate of unemployment will have to be tolerated.

The shape of the *PC* curve further suggests that when the unemployment rate is less than 5 per cent (that is, to the left of point *A*), the demand for labour is more than the supply and this tends to increase money wage rates. On the other hand, when the unemployment rate is more than 5½

per cent (to the right of point *A*), the supply of labour is more than the demand which tends to lower wage rates. The implication is that the wage rates will be stable at the unemployment rate *OA* which is equal to 5½ per cent per annum.

---

25. A. W. Phillips, "The Relation between Unemployment and the Rate of Change in Money Wage Rates in the United Kingdowm 1862-1956", *Economica*, November, 1958.

It is to be noted that *PC* is the "conventional" or original downward sloping Phillips curve which shows a stable and inverse relation betwen the rate of unemployment and the rate of change in wages.

Friedman's View : The Long-Run Phillips Curve

Economists have criticised and in certain cases modified the Phillips curve. They argue that the Phillips curve relates to the short run and it does not remain stable. It shifts with changes in expectations of inflation. In the long run, there is no trade-off between inflation and unemployment. These views have been expounded by Friedman and Phelps[27] in what has come to be known as the "accelerationist" or the "adaptive expectations" hypothesis.

According to Friedman, there is no need to assume a stable downward sloping Phillips curve to explain the trade-off between inflation and unemployment. In fact, this relation is a short-run phenomenon. But there are certain variables which cause the Phillips curve to shift over time and the most important of them is the expected rate of inflation. So long as there is discrepancy between the expected rate and the actual rate of inflation, the downward sloping Phillips curve will be found. But when this discrepancy is removed over the long run, the Phillips curve becomes vertical.

In order to explain this, Friedman introduces the concept of the *natural rate of unemployment*. In represents the rate of unemployment at which the economy normally settles because of its structural imperfections. It is the unemployment rate below which the inflation rate increases, and above which the inflation rate decreases. At this rate, there is neither a tendency for the inflation rate to increase or decrease. Thus the natural

rate of unemployment is defined as the rate of unemployment at which the actual rate of inflation equals the expected rate of inflation. It is thus an *equilibrium* rate of unemployment toward which the economy moves in the long run. In the long run, the Phillips curve is a vertical line at the natural rate of unemployment.

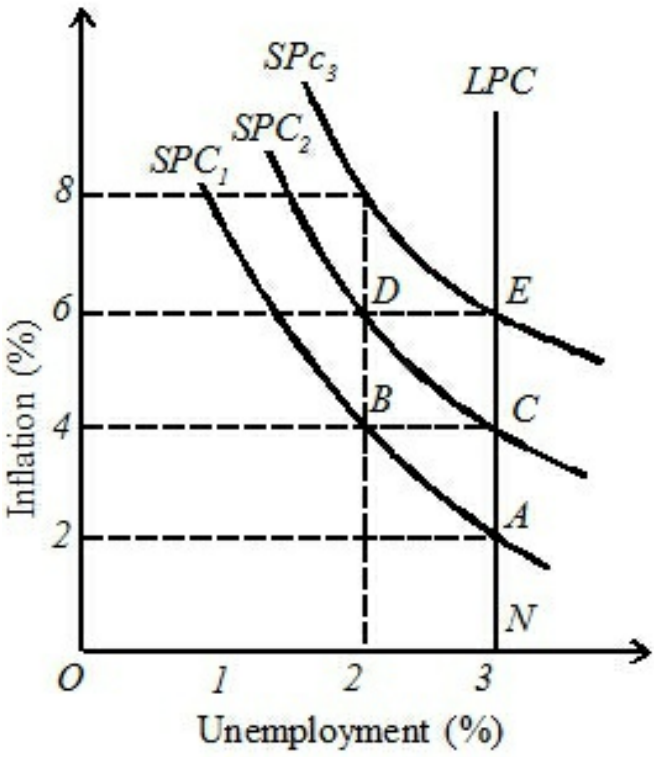


Fig. 11

This natural or equilibrium unemployment rate is not fixed for all times. Rather, it is determined by a number of structural characteristics of the labour and commodity markets within the economy. These may be minimum wage laws, inadequate employment information, deficiencies in manpower training, costs of labour mobility, and other market imperfections.But what causes the Phillips curve to shift over time is the expected rate of inflation. This refers to the extent the labour correctly forecasts inflation and can adjust wages to the forecast. Suppose the economy is experiencing a mild rate of inflation of 2 per cent and a natural rate of unemployment (*N*) of 3 per cent. At point *A* on the short-run Phillips curve SPC_1 in Figure 11, people expect this rate of inflation to continue in the future. Now assume that the government adopts a monetary-fiscal programme to raise aggregate demand in order to lower unemployment from 3 to 2 per cent. The increase in aggregate demand will raise the rate of inflation to 4 per cent consistent with the unemployment rate of 2 per cent. When the actual inflation rate (4 per cent) is greater than the expected inflation rate (2 per cent), the economy moves from point A to B along the SPC_1 curve and the unemployment rate temporarily falls to 2 per cent. This is achieved because the labour has been deceived. It expected the inflation rate of 2 per cent and based their

wage demands on this rate. But the workers eventually begin to realise that the actual rate of inflation is 4 per cent which now becomes their expected rate of inflation. Once this happens the short-run Phillips curve SPC_1 shifts to the right to SPC_2. Now workers demand increase in money wages to meet the higher expected rate of inflation of 4 per cent. They demand higher wages because they consider the present money wages to be inadequate in real terms. In other words, they want to keep up with higher prices and to eliminate fall in real wages. As a result, real labour costs will rise, firms will discharge workers and unemployment will rise from *B* (2%) to *C* (3%) with the shifting of the SPC_1 curve to SPC_2. At point *C*, the natural rate of unemployment is re-established at a higher rate of both the actual and expected inflation (4%).

---

26. R.G. Lipsey, "The Relationship between Unemployment and the Rate of change of money wage rates in the U.K. 1862-1957: A further analysis," *Economica,* Vol. 27, 1960.

* The difference between money wage rate and the rate of labour productivity is the inflation rate.

If the government is determined to maintain the level of unemployment at 2 per cent, it can do so only at the cost of higher rates of inflation. From point *C*, unemployment once again can be reduced to 2 per cent via increase in aggregate demand along the SPC_2 curve until we arrive at point *D*. With 2 per cent unemployment and 6 per cent inflation at point *D*, the expected rate of inflation for workers is 4 per cent. As soon as they adjust their expectations to the new situation of 6 per cent inflation, the short-run Phillips curve shifts up again to SPC_3, and the unemployment will rise back to its natural level of 3 per cent at point *E*. If points A, *C* and *E* are connected, they trace out a vertical long-run Phillips curve *LPC* at the natural rate of unemployment. On this curve, there is no trade-off between unemployment and inflation. Rather, any one of several rates of inflation at points *A*, *C* and *E* is compatible with the natural unemployment rate of 3 per cent. Any reduction in unemployment rate below its natural rate will be associated with an accelerating and ultimately explosive inflation. But this is only possible temporarily so long as workers overestimate or underestimate the inflation rate. In the long-run, the economy is bound to establish at the natural unemployment rate.

There is, therefore, no trade-off between unemployment and inflation except in the short run. This is because inflationary expectations are revised according to what has happened to inflation in the past. So when the actual rate of inflation, say, rises to 4 per cent in Figure 11, workers continue to expect 2 per cent inflation for a while and only in the long run they revise their expectations upward to 4 per cent. Since they adapt themselves to the expectations, it is called the *adaptive expections hypothesis*. According to this hypothesis, the expected rate of inflation always lags behind the actual rate. But if the actual rate remains constant, the expected rate would ultimately become equal to it. This leads to the conclusion that a short-run trade off exists between unemployment and inflation, but there is no long run trade-off between the two unless a continuously rising inflation rate is tolerated.

---

27. M. Friedman, "The Role of Monetary Policy," *A.E.R.*, March 1958 and E.S. Phelps, *Inflation Policy and Unemployment Theory*, 1973.

Its Criticisms

The accelerationist hypothesis of Friedman has been criticised on the following grounds :

1. The vertical long-run Phillips curve relates to steady rate of inflation. But this is not a correct view because the economy is always passing through a series of disequilibrium positions with little tendency to approach a steady state. In such a situation, expectations may be disappointed year after year.

2. Friedman does not give a new theory of how expectations are formed that would be free from theoretical and statistical bias. This makes his position unclear.

3. The vertical long-run Phillips curve implies that all expectations are satisfied and that people correctly anticipate the future inflation rates. Critics point out that people do not anticipate inflation rates correctly, particularly when some prices are almost certain to rise faster than others. There are bound to be disequilibria between supply and demand caused by uncertainty about the future and that is bound to increase the rate of

unemployment. Far from curing unemployment, a dose of inflation is likely to make it worse.

4. In one of his writings Friedman himself accepts the possibility that the long-run Phillips curve might not just be vertical, but could be positively sloped with increasing doses of inflation leading to increasing unemployment.

5. Some economists have argued that wage rates have not increased at a high rate of unemployment.

6. It is believed that workers have a money illusion. They are more concerned with the increase in their money wage rates than real wage rates.

7. Some economists regard the natural rate of unemployment as a mere abstraction because Friedman has not tried to define it in concrete terms.

8. Saul Hyman has estimated that the long-run Phillips curve is not vertical but is negatively sloped. According to Hyman, the unemployment rate can be permanently reduced if we are prepared to accept an increase in inflation rate.

Tobin's View

James Tobin[28] in his presidential address before the American Economic Association in 1971 proposed a compromise between the negatively sloping and vertical Phillips curves. Tobin believes that there is a Phillips curve within limits. But as the economy expands and employment grows, the curve becomes even more fragile and vanishes until it becomes vertical at some critically low rate of unemployment. Thus Tobin's

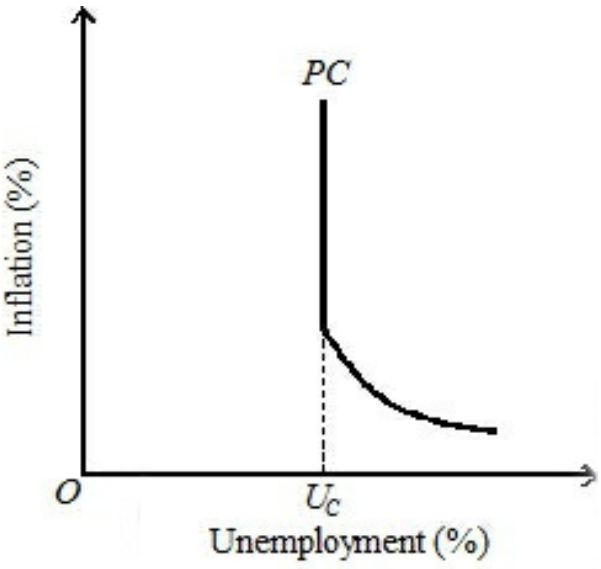


Fig. 12

Phillips curve is kinked-shaped, a part like a normal Phillips curve and the rest vertical, as shown in Figure 12. In the figure *Uc* is the critical rate of unemployment at which the Phillips curve becomes vertical where there is no trade-off between unemployment and inflation. According to Tobin, the vertical portion of the curve is not due to increase in the demand for more wages but emerges from imperfections of the labour market. At the *Uc* level, it is not possible to provide more employment because the job seekers have wrong skills or wrong age or sex or are in the wrong place. Regarding the normal portion of the Phillips curve which is negatively sloping, wages are sticky downward because labourers resist a decline in their relative wages. For Tobin, there is a wage-change floor in excess supply situations. In the range of relatively high unemployment to the right of *Uc* in the figure, as aggregate demand and inflation increase and involuntary unemployment is reduced, wage-floor markets gradually diminish. When all sectors of the labour market are above the wage floor, the level of critically low rate of unemployment Uc is reached.

Solow's View

Like Tobin, Robert Solow[29] does not believe that the Phillips curve is vertical at all rates of inflation. According to him, the curve is vertical at positive rates of inflation and is horizontal at negative rates of inflation, as shown in Figure 13. The basis of the Phillips curve *LPC* of the figure is that wages are sticky downward even in the face of heavy unemployment or deflation. But at a particular level of unemployment when the demand for labour increases, wages rise in the face of expected inflation. But since the Phillips curve *LPC* becomes vertical at that minimum level of unemployment, there is no trade-off between unemployment and inflation.

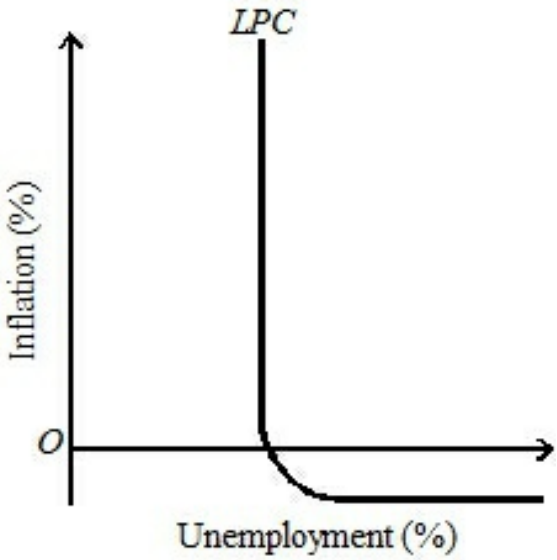


Fig. 13

Conclusion. The vertical Phillips curve has been accepted by the majority of economists. They agree that at unemployment rate of about 4 per cent, the Phillips curve becomes vertical and the trade-off between unemployment and inflation disappears. It is impossible to reduce unemployment below this level because of market imperfections.

RATIONAL EXPECTATIONS AND LONG-RUN PHILLIPS CURVE

In the Friedman-Phelps acceleration hypothesis of the Phillips curve, there is a short-run trade-off between unemployment and inflation but no long-run trade-off exists. The reason is that inflationary expectations are based on past behaviour of inflation which cannot be predicted accurately. Therefore, there is always an observed error so that the expected rate of inflation always lags behind the actual rate. But the expected rate of inflation is revised in accordance with the first period's experience of inflation by adding on some proportion of the observed error in the previous period so that the expected rate of inflation adjusts toward the actual rate.

Economists belonging to the rational expectations (Ratex) school have denied the possibility of any trade-off between inflation and unemployment even during the long run. According to them, the assumption implicit in Friedman's version that price expectations are formed mainly on the basis of the experience of past inflation is unrealistic. When people base their price expectations on this assumption, they are irrational. If they think like this during a period of rising prices, they will find that they were wrong. But rational people will not commit this mistake. Rather, they will use all available information to forecast future inflation more accurately.

---

28. "Inflation and Unemployment," *A.E.R.*, March 1972.

The rational expectations idea is explained in Figure 14 in relation to the Phillips curve. Suppose the unemployment rate is 3 per cent in the economy and the inflation rate is 2 per cent. We start at point *A* on the SPC_1 curve. In order to reduce unemployment, the government increases

the rate of money supply so as to stimulate the economy. Prices start rising. According to the Ratex hypothesis, firms have better information about prices in their own industry than about the general level of prices. They mistakenly think that the increase in prices is due to the increase in the demand for their products. As a result, they employ more workers in order to increase output. In this way, they reduce unemployment. The workers also mistake the rise in prices as related to their own industry. But wages rise as the demand for labour increases and workers think that the increase in money wages is an increase in real wages. Thus the economy moves upward on the short-run Phillips curve SPC_1 from point *A* to *B*. But soon workers and firms find that the increase in prices and wages is prevalent in most industries. Firms find that their costs have increased. Workers realise that their real wages have fallen due to the rise in the inflation rate to 4 per cent and they press for increase in wages. Thus the economy finds itself at the higher inflation rate due to government's monetary policy. As a result, it moves from point *B* to point *C* on the SPC_2 curve where the unemployment rate is 3 per cent which is the same before the government adopted an expansionary monetary policy.

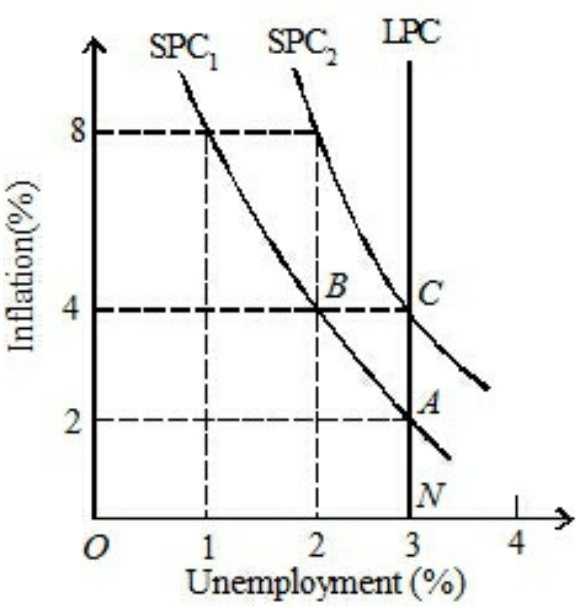


Fig. 14

When the government again tries to reduce unemployment by increasing the money supply, it cannot fool workers and firms who will now watch the movements of prices and costs in the economy. If firms expect higher costs with higher prices for their products, they are not likely to increase their production, as happened in the case of the SPC_1 curve. So far as workers are concerned, labour unions will demand higher wages to keep pace with prices moving up in the economy. When the government continues an expansionary monetary (or fiscal) policy, firms and workers get accustomed to it.They build their experience into their expectations. So when the government again adopts such a policy, firms raise prices of

their products to nullify the expected inflation so that there is no effect on production and employment. Similarly, workers demand higher wages in expectation of inflation and firms do not offer more jobs. In other words, firms and workers build expectations into their price policies and wage agreements so that there is no possibility for the actual rate of unempolyment to differ from the natural rate even during the short run.

---

29. "Down the Phillips Curve with Gun and Camera," in D.A. Belsley *et al.*, (eds), *Inflation, Trade and Taxes*, 1946.

POLICY IMPLICATIONS OF THE PHILLIPS CURVE

The Phillips curve has important policy implications. It suggests the extent to which monetary and fiscal policies can be used to control inflation without high levels of unemployment. In other words, it provides a guideline to the authorities about the rate of inflation which can be tolerated with a given level of unemployment. For this purpose, it is important to know the exact position of the Phillips curve. If the curve is PC_1 as in Figure 15,

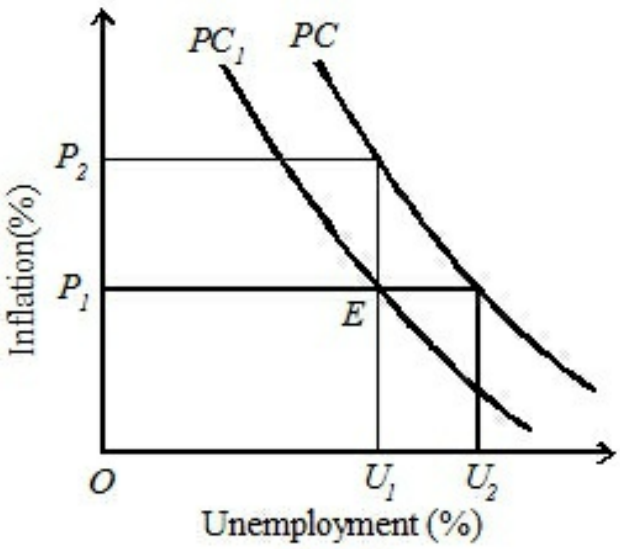


Fig. 15

where the labour productivity and the wage rate are equal at point *E*, both full employment and price stability would be possible. Again, a curve to the left of point *E* suggests full employment and price stability as consistent policy objectives. It implies that a lower level of inflation can be traded-off for a low level of unemployment. If, on the other hand, the Phillips curve is *PC* as in the figure, it suggests that the authorities will have to choose between price stability and more unemployment. Thus by observing the position of the Phillips curve, the authorities can decide about the nature of monetary and fiscal policies to be adopted. For instance, if the authorities find that the inflation rate P_2 is incompatible

with the unemployment rate U_1 of Figure 15, they would adopt such monetary and fiscal policies as to shift the Phillips curve *PC* to the left in the position of PC_1 curve. This will give a better trade-off between a lower inflation rate P_1 with the small level of unemployment U_1.

While explaining the natural rate of unemployment, Friedman pointed out that the only scope of public policy in influencing the level of unemployment lies in the short run in keeping with the position of the Phillips curve. He ruled out the possibility of influencing the long-run rate of unemployment because of the vertical Phillips curve.

According to him, the trade-off between unemployment and inflation does not exist and has never existed. However rapid the inflation might be, unemployment always tends to fall back to its natural rate which is not some irreducible minimum of unemployment. It can be lowered by removing obstacles in the labour market by reducing frictions. Therefore, public policy should improve the institutional structure to make the labour market responsive to changing patterns of demand. Moreover, some level of unemployment must be accepted as natural because of the existence of large number of part-time workers, unemployment compensation and other institutional factors.

Another implication is that unemployment is not a fitting aim for monetary expansion, according to Friedman. Therefore, employment above the natural rate can be reached at the cost of accelerating inflation, if monetary policy is adopted. In his words, “A little inflation will provide a boost at first—like a small dose of a drug for a new addict—but then it takes more and more inflation to provide the boost, just it takes a bigger and bigger dose of a drug to give a hardened addict a high.” Thus if the government wants to have a genuine full employment level at the natural rate, it must not use monetary policy to remove institutional restraints, restrictive practices, barriers to mobility, trade union coercion and similar obstacles to both the workers and the employers.

But economists do not agree with Friedman. They suggest that it is possible to reduce the natural rate of unemployment through labour market policies, whereby labour market can be made more efficient. So

the natural rate of unemployment can be reduced by shifting the long-run vertical Phillips curve to the left.

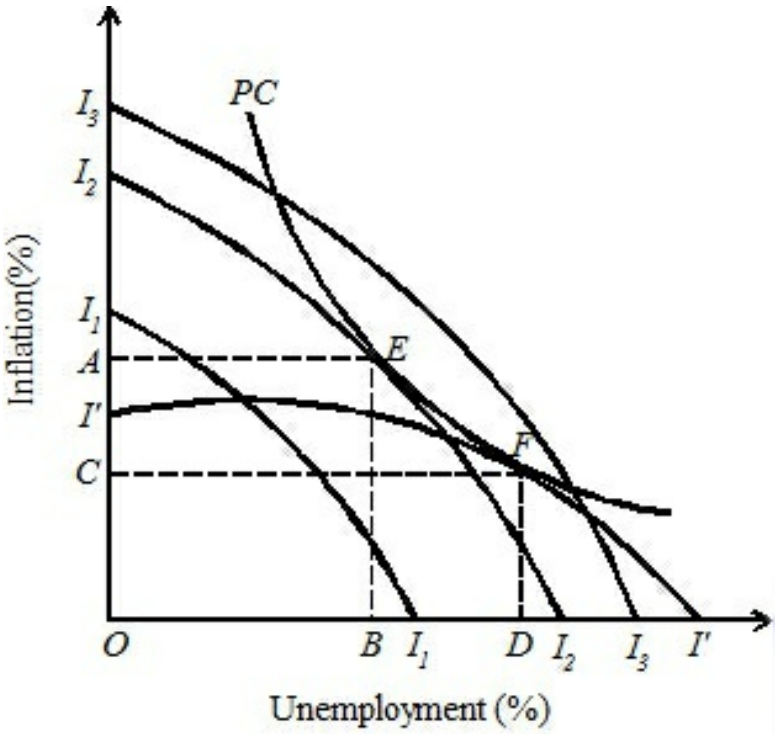


Fig. 16

But the policy implications of the Phillips curve are not so simple as they appear. The authorities are faced with certain constraints concerning the decision with regard to the rate of inflation that may be compatible with a particular rate of unemployment. Thus the problem of trade-off between inflation and unemployment is one of choice under constraints. This is illustrated in Figure 16. The constraints are a given Phillips curve *PC* and the indifference curves I_1 I_1, I_2 I_2, I_3 I_3 and *I"I"* representing the choice of authorities between unemployment and inflation. The indifference curves are concave to the origin because if the authorities want to reduce unemployment, they must have higher inflation and vice-versa. So they represent nagative utility. But the cuve I_2I_2 represents a higher level of public welfare than the curve I_1I_1, and the curve I_3I_3 still higher welfare than I_2I_2 curve. This is because any point on the lower curve represents a lower rate of unemployment and inflation than on a higher curve. The optimum trade-off point is *E* where the indifference curve I_1I_1 is tangent to the Phillips curve *PC* and where the trade-off is between *OA* rate of inflation and *OB* rate of unemployment. If, however, the public authorities adopt such monetary and fiscal policies whereby they want to have less inflation and more unemployment, the indifference curve becomes *1'1'*. This curve *I'I'* is tangent to the Phillips curve *PC* at *F* and the trade-off becomes *OC* of

inflation and *OD* of unemployment.

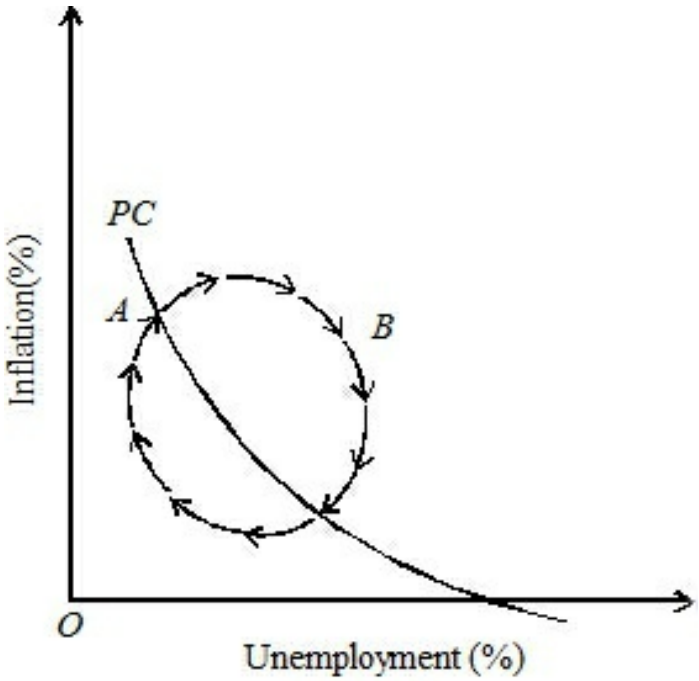


Fig. 17

It has been suggested by certain economists that there is a loop or orbit about the Phillips curve based on observed values of inflation and unemployment. This is illustrated in Figure 17. In the early expansion phase of the business cycle, the unemployment-inflation loop involves rising output with reduced inflation. This is due to demand-pull following an expansionary monetary or fiscal policy. In this phase of the cycle, the nomal relationship between inflation and unemployment suggested by the Phillips curve is maintained. It is shown by the movement of arrows at point *C* from below the *PC* curve when the rate of unemployment falls and the rate of inflation increases. If aggregate demand continues to increase, inflationary pressures gain momentum, and the dotted loop crosses the Phillips curve at point A. A tight monetary or fiscal policy will reduce aggregate demand. But the expectations of increase in prices will bring wage increases and inflation will be maintained at the previous rate. So unemployment will increase with no reduction in prices. This is revealed by the upper portion of the loop to the right of the Phillips curve. However, when excess demand is controlled and output increases, the rate of inflation starts falling from point B along with fall in the rate of unemployment. Thus we find that the conclusion of the Phillips curve holds in the early phase of the business cycle due to an expansionary monetary or fiscal policy. But in the downward phase the trade-off between inflation and unemployment goes contrary to the Phillips curve.

Johnson doubts about the applicability of the Phillips curve to the formulation of economic policy on two grounds. “On the one hand, the

curve represents only a statistical description of the mechanics of adjustment in the labour market, resting on a simple model of economic dynamics with little general and well-tested monetary theory behind it. On the other hand, it describes the behaviour of the labour market in a combination of periods of economic fluctuation and varying rates of inflation, conditions which presumably influenced the behaviour of the labour market itself, so that it may reasonably be doubted whether the curve would continue to hold its shape if an attempt were made by economic policy to pin the economy down to a point on it."[30]

STAGFLATION

Stagflation is a new term which has been added to economic literature in the 1970s. The word "stagflation" is the combination of stag *plus* flation, taking 'stag' from stagnation and 'flation' from inflation. Thus it is a paradoxical situation where the economy experiences stagnation or unemployment alongwith a high rate of inflation. It is, therefore, also called *inflationary recession*. The level of stagflation is measured in the US by the "discomfort index" which is a combination of the unemployment rate and the inflation rate measured by the price deflator for GNP.

One of the principal causes of stagflation has been restriction in the aggregate supply. When aggregate supply is reduced, there is a fall in output and employment and the price level rises. A reduction in aggregate supply may be due to a restriction in labour supply. The restriction in labour supply, in turn, may be caused by a rise in money wages on account of strong unions or by a rise in the legal minimum wage rate, or by increased tax rates which reduce work-effort on the part of workers.

When wages rise, firms are forced to reduce production and employment. Consequently, there is fall in real income and consumer expenditure. Since the decline in consumption will be less than the fall in real income, there will be excess demand in the commodity market which will push up the price level. The rise in the price level, in turn, reduces output and employment in the following three ways: (a) It reduces the real quantity of money, raises interest rates and brings a fall in investment expenditure. (b)

The rise in the price level reduces the real value of cash balances with the government and the private sector via the Pigou effect which reduces their consumption expenditure. (c) The rise in prices of domestic goods makes exports dearer for foreigners and make foreign goods relatively more attractive to domestic consumers, thereby adversely affecting domestic output and employment.

Another cause of restriction in aggregate supply is the increase in indirect taxes by central, state and local governments. When indirect taxes are increased, they raise costs and prices and reduce output and employment. Moreover, when the government increases taxes, it leads to the transfer of real purchasing power from the people to the government. As a result, aggregate demand falls, and output and employment are adversely affected. If, however, the government increases its expenditure equal to the increase in tax revenue, it would raise the price level further due to increase in additional demand.

Often, economies impose direct controls as a means of controlling inflation. But when such controls are removed, decontrolled sectors raise prices of their products with the result that wages rise and the wage-price spiral spreads to the entire economy. This, in turn, advesely affects production and employment through a decline in the real quantity of money, rise in interest rates, fall in investment via the Pigou effect, and exports becoming dearer and imports attractive. They contribute to stagflation.

Restriction on aggregate supply may also be caused by external factors such as rise in the world prices of foodgrains and crude oil prices. In all these cases, the domestic price level is raised by outside forces. When international prices of foodgrains and crude oil rise, they lead to the outflow of purchasing power away from domestic consumers. They accentuate inflation, raise wages and prices. As a result, the real quantity of money declines, interest rates rise and investment declines via the Pigou effect, making exports dearer and imports attractive, and domestic output and employment decline. They lead to stagflation.

The phenomenon of stagflation is illustrated in Fig. 18 where employment is measured on the horizontal axis and the price level on the vertical axis.

The initial equilibrium is at *E* where the demand curve *D* intersects the supply curve S and the price level is *OP* and the employment level is *ON*. When the aggregate supply is reduced due to any of the factors mentioned above, the supply curve *S* shifts to the left at S_1. The new equilibrium is at E_1 where S_1 intersects the *D* curve. Now the price level rises from *OP* to OP_1 and the level of employment declines from *ON* to ON_1.

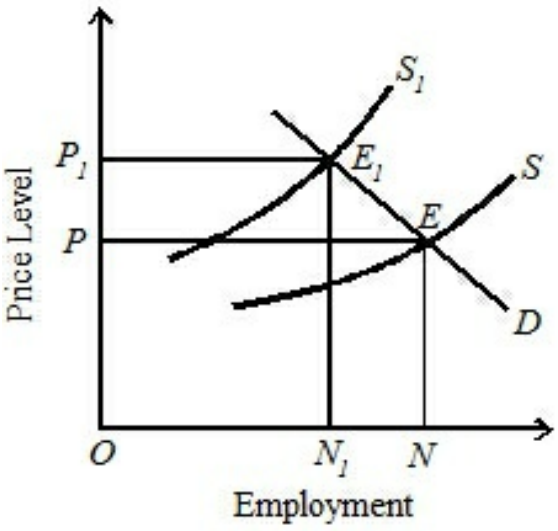


Fig. 18

---

30. *Op. cit.*, pp.132-3.

Measures to Control Stagflation

We have observed above that it is inflation that leads to stagflation. The US experience shows that if stagflation is controlled either by restrictive or expansionary measures, it will increase. Suppose restrictive demand managed monetary and fiscal measures are adopted, they tend to lower aggregate demand so that the new demand curve D_1 cuts the supply curve S_1 at point *E'* at the old price level *OP* in Fig 19. This policy reduces the level of employment further to *ON'* and at the same time lowers the price level from OP_1 to *OP*. Thus such a policy tends to increase unemployment by

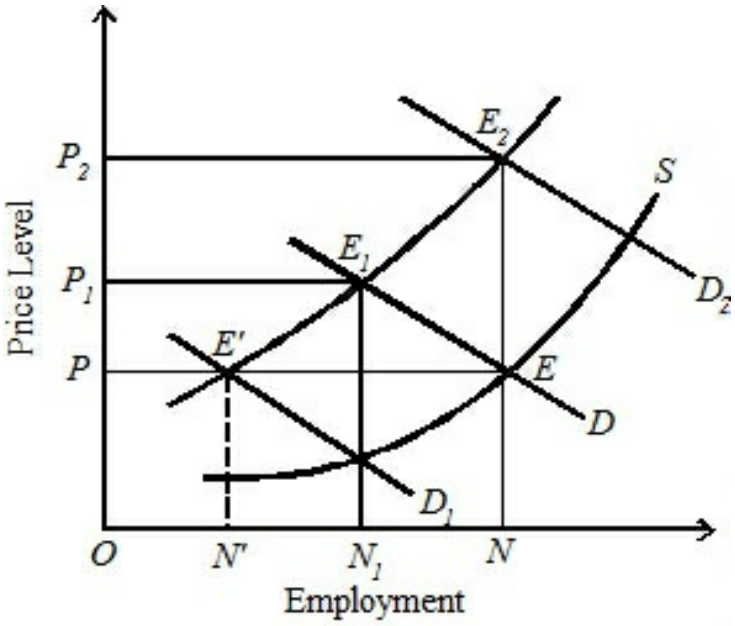


Fig. 19

N_1N' and reduces inflation by P_1P. Thus it fails to control stagflation. On the other hand, if expansionary demand managed monetary and fiscal policies are adopted, they will raise the aggregate demand so that the new demand curve D_2 cuts the supply curve S_1 at E_2 at the old employment level ON. This raises employment from ON_1 to ON but increases the price level to OP_2. Thus such a policy also fails to control stagflation because it generates more inflation combined with higher employment. Economists, therefore, suggest other measures which slow inflation and maintain higher employment.

First, minimum wages should not be raised at all.

Second, *tax-based income* policies should be started. These policies are different for individual and business firms. In the case of individuals, target rates of wage and price inflation are based on some reasonable economic forecast of inflation. Persons who accept wage increases below the target rates are rewarded with tax credits. Those who insist on wage increases above the target rates are levied a penalty tax. Similar is the case with business firms. Firms which keep wages down to the target rates are rewarded with a reduction in their business income tax. On the other hand, those who permit wage increases above the target rates are charged a penalty tax in addition to business income tax.

Third, there is need to introduce *income policies*. One of the important planks of the income policies is to link the increase of money wages to productivity increase. Thus the rate of increase of money wages should be limited to the overall rate of productivity increase. Further, prices should be reduced in those industries having above-average productivity growth. On the other hand, prices should be raised in industries where productivity is increasing less than the national average rate. Prices should be kept stable in industries where productivity is increasing at the national average rate. But such policies are difficult to implement in the case of an open country. If import prices of food and other consumer products rise, they tend to raise the domestic price level. This makes it difficult for unions to stick to wage agreements.

Fourth, the best policy measure is to reduce *personal and business taxes*

because they tend to reduce labour costs and raise demand for labour. Similarly, sales tax and excise duties should be reduced in order to prevent the price level from rising. To encourage state and local government to reduce state and local sales and excise taxes, the central government should sanction additional grants-in-aid to them.

Thus to combat stagflation, a vast spectrum of policy measures is needed.

CAUSES OF INFLATION

Inflation is caused when aggregate demand exceeds aggregate supply of goods and services. We analyse the factors which lead to increase in demand and shortage of supply.

Factors Affecting Demand

Both Keynesians and monetarists believe that inflation is caused by increase in aggregate demand. They point toward the following factors which raise it.

1. Increase in Money Supply. Inflation is caused by an increase in the supply of money which leads to increase in aggregate demand. The higher the growth rate of the nominal money supply, the higher is the rate of inflation. Modern quantity theorists do not believe that true inflation starts after the full employment level. This view is realistic because all advanced countries are faced with high levels of unemployment and high rates of inflation.

2. Increase in Disposable Income. When the disposable income of the people increases, it raises their demand for goods and services. Disposable income may increase with the rise in national income or reduction in taxes or reduction in the saving of the people.

3. Increase in Public Expenditure. Government activities have been expanding much with the result that government expenditure has also been increasing at a phenomenal rate, thereby raising aggregate demand for goods and services. Governments of both developed and developing

countries are providing more facilities under public utilities and social services, and also nationalising industries and starting public enterprises with the result that they help in increasing aggregate demand.

4. Increase in Consumer Spending. The demand for goods and services increases when consumer expenditure increases. Consumers may spend more due to conspicuous consumption or demonstration effect. They may also spend more when they are given credit facilities to buy goods on hire-purchase and instalment basis.

5. Cheap Monetary Policy. Cheap monetary policy or the policy of credit expansion also leads to increase in the money supply which raises the demand for goods and services in the economy. When credit expands, it raises the money income of the borrowers which, in turn, raises aggregate demand relative to supply, thereby leading to inflation. This is also known as *credit-induced inflation.*

6. Deficit Financing. In order to meet its mounting expenses, the government resorts to deficit financing by borrowing from the public and even by printing more notes. This raises aggregate demand in relation to aggregate supply, thereby leading to inflationary rise in prices. This is also known as *deficit-induced* inflation.

7. Expansion of the Private Sector. The expansion of the private sector also tends to raise the aggregate demand. For huge investments increase employment and income, thereby creating more demand for goods and services. But it takes time for the output to enter the market. This leads to rise in prices.

8. Black Money. The existence of black money in all countries due to corruption, tax evasion etc. increases the aggregate demand. People spend such unearned money extravagantly, thereby creating unnecessary demand for commodities. This tends to raise the price level further.

9. Repayment of Public Debt. Whenever the government repays its past internal debt to the public, it leads to increase in the money supply with the public. This tends to raise the aggregate demand for goods and services and to rise in prices.

10. Increase in Exports. When the demand for domestically produced goods increases in foreign countries, this raises the earnings of industries producing export commodities. These, in turn, create more demand for goods and services within the economy, thereby leading to rise in the price level.

Factors Affecting Supply

There are also certain factors which operate on the opposite side and tend to reduce the aggregate supply. Some of the factors are as follows :

1. Shortage of Factors of Production. One of the important causes affecting the supplies of goods is the shortage of such factors as labour, raw materials, power supply, capital, etc. They lead to excess capacity and reduction in industrial production, thereby raising prices.

2. Industrial Disputes. In countries where trade unions are powerful, they also help in curtailing production. Trade unions resort to strikes and if they happen to be unreasonable from the employers' viewpoint and are prolonged, they force the employers to declare lock-outs. In both cases. industiral production falls, thereby reducing supplies of goods. If the unions succeed in raising money wages of their members to a very high level than the productivity of labour, this also tends to reduce production and supplies of goods. Thus they tend to raise prices.

3. Natural Calamities. Drought or floods is a factor which adversely affects the supplies of agricultural products. The latter, in turn, create shortages of food products and raw materials, thereby helping inflationary pressures.

4. Artificial Scarcities. Artificial scarcities are created by hoarders and speculators who indulge in black marketing. Thus they are instrumental in reducing supplies of goods and raising their prices.

5. Increase in Exports. When the country produces more goods for export than for domestic consumption, this creates shortages of goods in the domestic market. This leads to inflation in the economy.

6. Lop-sided Production. If the stress is on the production of comfort, luxury, or basic products to the neglect of essential consumer goods in the country, this creates shortages or consumer goods. This again causes inflation.

7. Law of Diminishing Returns. If industries in the country are using old machines and outmoded methods of production, the law of diminishing returns operates. This raises cost per unit of production, thereby raising the prices of products.

8. International Factors. In modern times, inflation is a worldwide phenomenon. When prices rise in major industrial countries, their effects spread to almost all countries with which they have trade relations. Often the rise in the price of a basic raw material like petrol in the international market leads to rise in the prices of all related commodities in a country.

MEASURES TO CONTROL INFLATION

We have studied above that inflation is caused by the failure of aggregate supply to equal the increase in aggregate demand. Inflation can, therefore, be controlled by increasing the supplies of goods and services and reducing money incomes in order to control aggregate demand. The various methods are usually grouped under three heads: monetary meausres, fiscal measures and other measures.

Monetary Measures

Monetary measures aim at reducing money incomes.

(a) Credit Control. One of the important monetary measures is *monetary policy*. The central bank of the country adopts a number of methods to control the quantity and quality of credit. For this purpose, it raises the bank rates, sells securities in the open market, raises the reserve ratio, and adopts a number of selective credit control measures, such as raising margin requirements and regulating consumer credit.

Monetary policy may not be effective in controlling inflation, if inflation

is due to cost-push factors. Monetary policy can only be helpful in controlling inflation due to demand-pull factors.

(b) Demonetisation of Currency. However, one of the monetary measures is to demonetise currency of higher denominations. Such a measures is usually adopted when there is abundance of black money in the country.

(c) Issue of New Currency. The most extreme monetary measure is the issue of new currency in place of the old currency. Under this system, one new note is exchanged for a number of notes of the old currency. The value of bank deposits is also fixed accordingly. Such a measure is adopted when there is an excessive issue of notes and there is hyperinflation in the country. It is a very effective measure. But is inequitable for its hurts the small depositors the most.

Fiscal Measures

Monetary policy alone is incapable of controlling inflation. It should, therefore, be supplemented by fiscal measures. Fiscal measures are highly effective for controlling government expenditure, personal consumption expenditure, and private and public investment. The principal fiscal measures are the following :

(a) Reduction in Unnecessary Expenditure. The government should reduce unnecessary expenditure on non-development activities in order to curb inflation. This will also put a check on private expenditure which is dependent upon government demand for goods and services. But it is not easy to cut government expenditure. Though this measure is always welcome but it becomes difficult to distinguish between essential and non-essential expenditure. Therefore, this measure should be supplemented by taxation.

(b) Increase in Taxes. To cut personal consumption expenditure, the rates of personal, corporate and commodity taxes should be raised and even new taxes should be levied, but the rates of taxes should not be so high as to discourage saving, investment and production. Rather, the tax system should provide larger incentives to those who save, invest and produce

more. Further, to bring more revenue into the tax-net, the government should penalise the tax evaders by imposing heavy fines. Such measures are bound to be effective in controlling inflation. To increase the supply of goods within the country, the government should reduce import duties and increase export duties.

(c) Increase in Savings. Another measure is to increase savings on the part of the people. This will tend to reduce disposable income with the people, and hence personal consumption expenditure. But due to the rising cost of living, people are not in a position to save much voluntarily. Keynes, therefore, advocated compulsory savings or what he called 'deferred payment' where the saver gets his money back after some years. For this purpose, the government should float public loans carrying high rates of interest, start saving schemes with prize money, or lottery for long periods, etc. It should also introduce compulsory provident fund, provident fund-cum-pension schemes, etc. All such measures increase savings and are likely to be effective in controlling inflation.

(d) Surplus Budgets. An important measure is to adopt anti-inflationary budgetary policy. For this purpose, the government should give up deficit financing and instead have surplus budgets. It means collecting more in revenues and spending less.

(e) Public Debt. At the same time, it should stop repayment of public debt and postpone it to some future date till inflationary pressures are controlled within the economy. Instead, the government should borrow more to reduce money supply with the public.

Like monetary measures, fiscal measures alone cannot help in controlling inflation. They should be supplemented by monetary, non-monetary and non-fiscal measures.

Other Measures

The other types of measures are those which aim at increasing aggregate supply and reducing aggregate demand directly.

(a) To Increase Production. The following measures should be adopted

to increase production: (i) One of the foremost measures to control inflation is to increase the production of essential consumer goods like food, clothing, kerosene oil, sugar, vegetable oils, etc. (ii) If there is need, raw materials for such products may be imported on preferential basis to increase the production of essential commodities. (iii) Efforts should also be made to increase productivity. For this purpose, industrial peace should be maintained through agreements with trade unions, binding them not to resort to strikes for some time. (iv) The policy of rationalisation of industries should be adopted as a long-term measure. Rationalisation increases productivity and production of industries through the use of brain, brawn and bullion. (v) All possible help in the form of latest technology, raw materials, financial help, subsidies, etc. shold be provided to different consumer goods sectors to increase production.

(b) Rational Wage Policy. Another important measure is to adopt a rational wage and income policy. Under hyperinflation, there is a wage-price spiral. To control this, the government should freeze wages, incomes, profits, dividends, bonus, etc. But such a drastic measure can only be adopted for a short period as it is likely to antagonise both workers and industrialists. Therefore, the best course is to link increase in wages to increase in productivity. This will have a dual effect. It will control wages and at the same time increase productivity, and hence raise production of goods in the economy.

(c) Price Control. Price control and rationing is another measure of direct control to check inflation. Price control means fixing an upper limit for the prices of essential consumer goods. They are the maximum prices fixed by law and anybody charging more than these prices is punished by law. But it is difficult to administer price control.

(d) Rationing. Rationing aims at distributing consumption of scarce goods so as to make them available to a large number of consumers. It is applied to essential consumer goods such as wheat, rice, sugar, kerosene oil, etc. It is meant to stabilise the prices of necessaries and assure distributive justice. But it is very inconvenient for consumers because it leads to queues, artificial shortages, corruption and black marketing. Keynes did not favour rationing for it “involves a great deal of waste, both

of resources and of employment."

Conclusion. From the various monetary, fiscal and other measures discussed above, it becomes clear that to control inflation, the government should adopt all measures simultaneously. Inflation is like a hydra-headed monster which should be fought by using all the weapons at the command of the government.

EFFECTS OF INFLATION*

Inflation affects different people differently. This is because of the fall in the value of money. When price rises or the value of money falls, some groups of the society gain, some lose and some stand in-between. Broadly speaking. there are two economic groups in every society, the fixed income group and the flexible income group. People belonging to the first group lose and those belonging to the second group gain. The reason is that the price movements in the case of different goods, services, assets, etc. are not uniform. When there is inflation, most prices are rising, but the rates of increase of individual prices differ much. Prices of some goods and services rise faster, of others slowly and of still others remain unchanged. We discuss below the effects of inflation on redistribution of income and wealth, production, and on the society as a whole.

1. Effects on Redistribution of Income and Wealth

There are two ways to measure the effects of inflation on the redistribution of income and wealth in a society. *First*, on the basis of the change in the real value of such factor incomes as wages, salaries, rents, interest, dividends and profits. *Second*, on the basis of the size distribution of income over time as a result of inflation. i.e. whether the incomes of the rich have increased and that of the middle and poor classes have declined with inflation. Inflation brings about shifts in the distribution of real income from those whose money incomes are relatively inflexible to those whose money incomes are relatively flexible.

---

* It is also known as effects of changes in the value of money.

The poor and middle classes suffer because their wages and salaries are more or less fixed but the prices of commodities continue to rise. They become more impoverished. On the other hand, businessmen, industrialists, traders, real estate holders, speculators, and others with variable incomes gain during rising prices. The latter category of persons become rich at the cost of the former group. There is unjustified transfer of income and wealth from the poor to the rich. As a result, the rich roll in wealth and indulge in conspicuous consumption, while the poor and middle classes live in abject misery and poverty.

But which income group of society gains or loses from inflation depends on who anticipates inflation and who does not. Those who correctly anticipate inflation, they can adjust their present earnings, buying, borrowing, and lending activities against the loss of income and wealth due to inflation. They, therefore, do nor get hurt by the inflation. Failure to anticipate inflation correctly leads to redistribution of income and wealth. In practice, all persons are unable to anticipate and predict the rate of inflation correctly so that they cannot adjust their economic behaviour accordingly. As a result, some persons gain while others lose. The net result is redistibution of income and wealth. The effects of inflation on different groups of society are discussed below.

(1) Debtors and Creditors. During periods of rising prices, debtors gain and creditors lose. When prices rise, the value of money falls. Though debtors return the same amount of money, but they pay less in terms of goods and services. This is because the value of money is less than when they borrowed the money. Thus the burden of the debt is reduced and debtors gain. On the other hand, creditors lose. Although they get back the same amount of money which they lent, they receive less in real terms because the value of money falls. Thus inflation brings about a redistribution of real wealth in favour of debtors at the cost of creditors.

(2) Salaried Persons. Salaried workers such as clerks, teachers, and other white collar persons lose when there is inflation. The reason is that their salaries are slow to adjust when prices are rising.

(3) Wage Earners. Wage earners may gain or lose depending upon the speed with which their wages adjust to rising prices. If their unions are

strong, they may get their wages linked to the cost of living index. In this way, they may be able to protect themselves from the bad effects of inflation. But the problem is that there is often a time lag between the raising of wages by employees and the rise in prices. So workers lose because by the time wages are raised, the cost of living index may have increased further. But where the unions have entered into contractual wages for a fixed period, the workers lose when prices continue to rise during the period of contract. On the whole, the wage earners are in the same position as the white collar persons.

(4) Fixed Income Group. The recipients of transfer payments such as pensions, unemployment insurance, social security, etc. and recipients of interest and rent live on fixed incomes. Pensioners get fixed pensions. Similarly the rentier class consisting of interest and rent receivers get fixed payments. The same is the case with the holders of fixed interest bearing securities, debentures and deposits. All such persons lose because they receive fixed payments, while the value of money continues to fall with rising prices. Among these groups, the recipients of transfer payments belong to the lower income group and the rentier class to the upper income group. Inflation redistributes income from these two groups toward the middle income group comprising traders and businessmen.

(5) Equity Holders or Investors. Persons who hold shares or stocks of companies gain during inflation. For when prices are rising, business activities expand which increase profits of companies. As profits increase, dividends on equities also increase at a faster rate than prices. But those who invest in debentures, securities, bonds, etc. which carry a fixed interest rate lose during inflation because they receive a fixed sum while the purchasing power is falling.

(6) Businessmen. Businessmen of all types, such as producers, traders and real estate holders gain during periods of rising prices. Take producers first. When prices are rising, the value of their inventories (goods in stock) rise in the same proportion. So they profit more when they sell their stored commodities. The same is the case with traders in the short run. But producers profit more in another way. Their costs do not rise to the extent of the rise in the prices of their goods. This is because prices of raw

materials and other inputs and wages do not rise immediately to the level of the price rise. The holders of real estates also profit during inflation because the prices of landed property increase much faster than the general price level.

(7) Agriculturists. Agriculturists are of three types, landlords, peasant proprietors, and landless agricultural workers. Landlords lose during rising prices becaused they get fixed rents. But peasant proprietors who own and cultivate their farms gain. Prices of farm products increase more than the cost of production. For prices of inputs and land revenue do not rise to the same extent as the rise in the prices of farm products. On the other hand, the landless agricultural workers are hit hard by rising prices. Their wages are not raised by the farm owners, because trade unionism is absent among them. But the prices of consumer goods rise rapidly. So landless agricultural workers are losers.

(8) Government. The government as a debtor gains at the expense of households who are its prinicipal creditors. This is because interest rates on government bonds are fixed and are not raised to offset expected rise in prices. The government, in turn, levies less taxes to service and retire its debt. With inflation, even the real value of taxes is reduced. Thus redistribution of wealth in favour of the government accrues as a benefit to the tax-payers. Since the tax-payers of the government are high-income groups, they are also the creditors of the government because it is they who hold government bonds. As creditors, the real value of their assets decline and as tax-payers, the real value of their liabilities also declines during inflation. The extent to which they will be gainers or losers on the whole is a very complicated calculation.

Conclusion. Thus inflation redistributes income from wage earners and fixed income groups to profit recipients, and from creditors to debtors. So far as wealth redistibutions are concerned, the very poor and the very rich are more likely to lose than middle income groups. This is because the poor hold what little wealth they have in monetary form and have few debts, whereas the very rich hold a substantial part of their wealth in bonds and have relatively few debts. On the other hand, the middle income groups are likely to be heavily in debt and hold some wealth in

common stocks as well as in real assets.

2. Effects on Production

When prices start rising production is encouraged. Producers earn windfall profits in the future. They invest more in anticipation of higher profits in the future. This tends to increase employment, production and income. But this is only possible up to the full employment level. Further increase in investment beyond this level will lead to severe inflationary pressures within the economy because prices rise more than production as the resources are fully employed. So inflation adversely affects production after the level of full employment. The adverse effects of inflation on production are discussed below.

(1) Misallocation of Resources. Inflation causes misallocation of resources when producers divert resources from the production of essential to non-essential goods from which they expect higher profits.

(2) Changes in the System of Transactions. Inflation leads to changes in transactions pattern of producers. They hold a smaller stock of real money holdings against unexpected contingencies than before. They devote more time and attention to converting money into inventories or other financial or real assets. It means that time and energy are diverted from the production of goods and services and some resources are used wastefully.

(3) Reduction in Production. Inflation adversely affects the volume of production because the expectation of rising prices alongwith rising costs of inputs bring uncertainty. This reduces production.

(4) Fall in Quality. Continuous rise in prices creates a seller's market. In such a situation, producers produce and sell sub-standard commodities in order to earn higher profits. They also indulge in adulteration of commodities.

(5) Hoarding and Blackmarketing. To profit more from rising prices, producers hoard stocks of their commodities. Consequently, an artificial scarcity of commodities is created in the market. Then the producers sell their products in the black market which increases inflationary pressures.

(6) Reduction in Saving. When prices rise rapidly, the propensity to save declines because more money is needed to buy goods and services than before. Reduced saving adversely affects investment and capital formation. As a result, production is hindered.

(7) Hinders Foreign Capital. Inflation hinders the inflow of foreign capital because the rising costs of materials and other inputs make foreign investment less profitable.

(8) Encourages Speculation. Rapidly rising prices create uncertainty among producers who indulge in speculative activities in order to make quick profits. Instead of engaging themselves in productive activities, they speculate in various types of raw materials required in production.

3. Other Effects

Inflation leads to a number of other effects which are discussed as under:

(1) Government. Inflation affects the government in various ways. It helps the government in financing its activities through inflationary finance. As the money incomes of the people increase, government collects that in the form of taxes on incomes and commodities. So the revenues of the government increase during rising prices. Moreover, the real burden of the public debt decreases when prices are rising. But the government expenses also increase with rising production costs of public projects and enterprises and increase in administrative expenses as prices and wages rise. On the whole, the government gains under inflation because rising wages and profits spread an illusion of prosperity within the country.

(2) Balance of Payments. Inflation involves the sacrificing of the advantages of international specialisation and division of labour. It affects adversely the balance of payments of a country. When prices rise more rapidly in the home country than in foreign countries, domestic products become costlier compared to foreign products. This tends to increase imports and reduce exports, thereby making the balance of payments unfavourable for the country. This happens only when the country follows a fixed exchange rate policy. But there is no adverse impact on the

balance of payments if the country is on the flexible exchange rate system.

(3) Exchange Rate. When prices rise more rapidly in the home country than in foreign countries, it lowers the exchange rate in relation to foreign currencies.

(4) Collapse of the Monetary System. If hyperinflation persists and the value of money countinues to fall many times in a day, it ultimately leads to the collapse of the monetary system, as happened in Germany after World War 1.

(5) Social. Inflation is socially harmful. By widening the gulf between the rich and the poor, rising prices create discontentment among the masses. Pressed by the rising cost of living, workers resort to strikes which lead to loss in production. Lured by profits, people resort to hoarding, blackmarketing, adulteration, manufacture of substandard commodities, speculation, etc. Corruption spreads in every walk of life. All this reduces the efficiency of the economy.

(6) Political. Rising prices also encourage agitations and protests by political parties opposed to the government. And if they gather momentum and become unhandy they may bring the downfall of the government. Many governments have been sacrificed at the alter of inflation.

INFLATION AS A TAX

Inflation operates like a tax when redistribution results in goods and services being transferred to the government from the poeple. If falls heavily on those least able to pay. When the government issues more money to finance its budget deficit, to repay its past debt and to meet its rising demand for goods and services during inflation, it acts as a tax on the people and it transfers purchasing power to the govenment. High inflation rates decrease the purchasing power of money with the poeple and discourage them from holding money. The rate of inflation is the rate of inflation tax. The inflation tax is defined as the decline in purchasing power of money due to inflation. It is calculated as :

$$M \times \frac{i}{(1+i)}$$

Where *M* is the average money at year-ending and year-beginnning and *i* is the decimal inflation rate measured by the change in consumer price index (*CPI*). The formula tells that the period for which prices rise by *i*, each money unit loses *i*/ (1+*i*) of its purchasing power.

Inflation as a tax on holding real money balances is explained in terms of Figure 20, where the level of real money balances is measured on the horizontal axis and the interest rate on the vertical axis. Suppose the government issues money to finance its budget deficit which will raise the price level and cause the real money stock to fall. Assuming that the initial price level is stable, and the level of real income is constant, the money interest rate (*i*) is equal to the real interest rate (*r*). We begin the analysis by further assuming zero expected rate of inflation which is equal to the money interest rate (*i*). In such as economy, the demand for real money balances as a function of the money interest rate is shown by the curve LL_1. If the money interest rate consistent with the stable price level is *i*, the amount of real money balances people wish to hold is (*M*/*P*). If the expected rate of inflation is *E*%, the interest rate rises to r_1 and the level of real cash balances falls from (*M*/*P*) to $(M/P)_1$. This means that as soon as the government announces the expected rate of inflation to be *E*% (i-r_1), everybody desiring to reduce his real cash balances will purchase physical assets and stocks of consumer and other goods, and the price level will rise in the

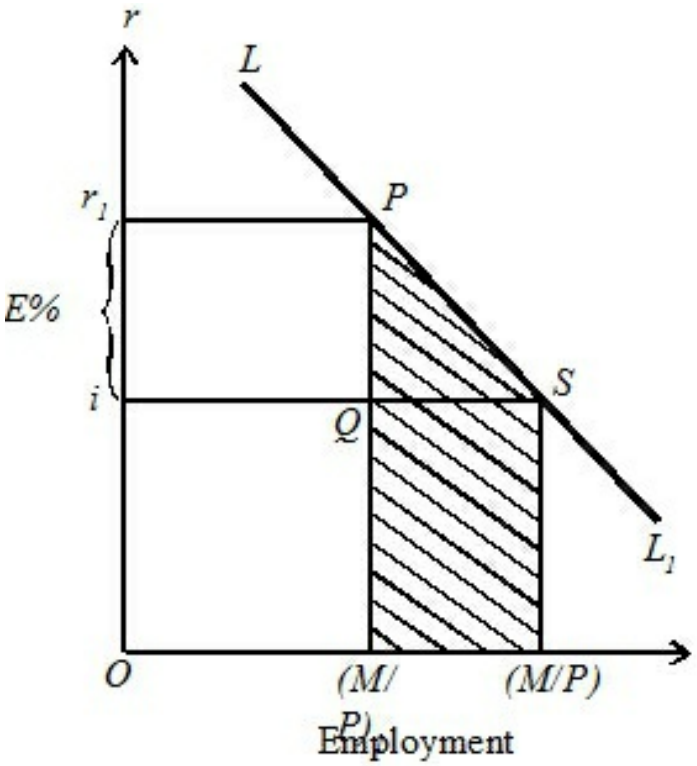


Fig. 20

proportion $(M/P)/(M/P)_1$. The proceeds of the tax in real terms are equal to the rectangle r_1 PQ i which is the inflation tax revenue to the government. The tax base is the amount of real money held by the public which is $(M/P)_1$ $(= iQ)$, and the tax rate is the inflation rate $(i–r_1)$

As a result of high inflation rate, the asset holders pay the inflation tax by losing purchasing power on their money holdings. The government as the issuer of money collects the tax in the form of a reduction in the real value of its liabilities. When the government pays interest on these liabilities, it returns some of the tax to holders of money. In practice, central banks do not pay so much interest as to offset the tax on the money issued by them. They pay no interest on currency and usually pay an interest rate on reserves below the market rate.

COSTS OF INFLATION

The costs of inflation may be economic or social loss arising from the effects of inflation. Assuming that people hold only non-interest bearing money in the form of currency issued by the government and demand deposits of banks, the costs of inflation refer to the loss in real money balances held by individuals and businesses. Since money does not bear a rate of interest, the opportunity cost of holding money rises with the inflation rate which, in turn, reduces the demand for real money balances.

Individuals and business enterprises hold cash balances because they yield utility to them. At a higher rate of inflation, they find the purchasing power of the money balances diminishing. In other words, they find that they require more real money balances than before when there is inflation. The costs of inflation arise when they try to change their existing system of transactions or payments to adjust to a smaller stock of real cash holdings. Individuals or households visit the markets more frequently to buy goods. Business enterprises visit banks more often, increase the frequency of ordering inventories, devote more time and attention in converting money into inventories or financial and real assets. Thus the change in the transactions or payments patterns of individuals and business enterprises require more time and energy then before. It leads to

the diversion of resources from productive to unproductive uses when they are required to visit markets and banks more frequently, maintain excessive inventories of consumer and producer goods, etc. When the real money balances with the people are reduced due to higher expected rate of inflation, their peace of mind is also disturbed. Thus "the ultimate social cost of anticipated inflation is the wasteful use of resources to economise holdings of currency and other non-interest bearing means of payment."

Another social cost of inflation is in terms of the Phillips curve analysis. When inflation starts and is expected to continue, any attmept to reduce its rate of increase will lead to more unemployment. Increase in unemployment is a loss to the economy in terms of the goods and services which cannot be produced because people available for employment are not used.

The majority of economists also regard the redistributive effects of inflation as the cost of inflation.*

The social cost of inflation can be measured in terms of Figure 20. The curve LL_1 is the demand curve for real cash balances which can be interpreted as the *MP* (utility) curve of real cash balances. When the rate of inflation is zero, the real interest rate is equal to the money interest rate at *i*. The demand for real cash balanaces is (*M*/*P*). The area under the demand curve LL_1 over a given segment of the horizontal axis measures the flow of productivity (utility) from the indicated quantity of real money balances. When inflation increases at the expected rate of *E*% ($i–r_1$) the interest rate rises from *i* to r_1 and the demand for real cash balances falls to $(M/P)_1$. This reduction in real cash balances by $(M/P)–(M/P)_1$ is the social cost of inflation which is measured by the shaded area $(M/P)_1$ *PS* (*M*/*P*). This area "measures the aggregate loss of productivity (utility) resulting from the destruction of real cash balances which occurs when prices rise initially at the announcement that there will be an inflation. The further rise of prices representing the inflation itself is merely sufficient to keep real balances at their new low level and so guarantee that this loss of productivity (utility) will continue as long as the inflation does."

DEFLATION

The opposite of inflation is deflation. It is "a state in which the value of money is rising i.e. prices are falling." It is usually associated with falling activity and employment. As pointed out by Coulborn, "Involuntary unemployment is the hall-mark of deflation." Deflation is caused when prices are falling more than proportionately to the output of goods and services in the economy as a result of decrease in the money supply.

---

* May be discussed briefly from the previous section.

Sometimes, deflation is confused with *disinflation*. Deflation is a situation when prices fall along with reduction in output and employment. Disinflation, on the other hand, is a situation when prices are reduced deliberately but output and employment remain unaffected. According to Coulborn, "A lowering of prices, income, and expenditures, when they would be beneficial, would be disinflation."

Effects of Deflation

The effects of deflation are the reverse of inflaton.** Deflation affects different groups differently. Persons with fixed incomes such as workers, white collar salaried workers, pensioners, the rentier class, etc. gain because the value of money rises with falling prices. On the other hand, all types of producers such as industrialists and farmers lose with falling prices. Traders and equity holders also lose. Thus deflation affects adversely the distribution of income and wealth. When prices are falling, the purchasing power is increasing. So the lower, middle, and other classes with low incomes gain. On the other hand, businessmen, industrialists, traders, real estate holders, and others with variable incomes are hit hard and their profits decline with deflation. But this does not mean that there is improvement in income distribution. Rather, the low income groups suffer more because of the fall in employment and income. So both the better off and the worse off feel discontented under deflation.

Deflation also affects production adversely. With falling prices, production falls because income and employment are also declining and

the aggregate demand is on the decline. Commodities start accumulating. Profits fall. Small firms close down. Unemployment spreads. This vicious cricle of fall in demand, production, employment, income and aggregate demand leads to a depression.

The government also suffer under deflation. Revenues from direct and indirect taxes decline. The real burden of public debt increases. Development of the economy suffers because the government is unable to increase public expenditure.

COMPARISON BETWEN INFLATION AND DEFLATION

Of the two—inflation and deflation—which is better than the other. Of course, both are equally bad in their effects on the society. But inflation is the lesser evil. As pointed out by Keynes, “Inflation is unjust, deflation is inexpedient. Of the two deflation is worse. Inflation brings about rising prices and redistribution of income in favour of the better-off classes. On the other hand, deflation leads to fall in output, employment and income. Of all the evils in a capitalist society, unemployment leading to poverty is the worst. We discuss below why Keynes regarded inflation as unjust and deflation inexpedient.

** For a detailed account write just on the basis of arguments opposite to that given under "Effects of Inflation".

Inflation is unjust because it widens the gulf between the rich and the poor. It makes the rich richer at the cost of the poor. On the other hand, the poor are made poorer. The poor and low income classes suffer because their wages and salaries do not rise to the extent prices rise. It becomes difficult for them to make both ends meet with rising prices of consumer goods. On the other hand, businessmen, traders, industrialists, real estate holders, speculators, etc. gain because their profits and incomes increase much more than the rise in prices. So they are not affected by the fall in purchasing power when prices are rising. Thus it leads to inequalities of income and wealth.

When the government resorts to deficit financing to meet its rising

expenditure during inflationary pressures, it increases the demand for goods and services. This deprives the people of the use of essential goods, thereby creating shortages and hardships for the common man.

Again, inflation is unjust because persons who save are losers in the long run. When prices are rising, the value of money is falling. Since savers are mostly the low and middle income groups who save for a varietyof reasons, they are the losers. Their savings lying in deposits are reduced automatically in real terms as inflationary pressures increase.

Inflation is unjust because it is socially harmful. People are lured to amass wealth by unscrupulous means. They, therefore, resort to hoarding, black-marketing, adulteration, manufature of sub-standard commodities, speculation, etc. Corruption spreads in every walk of life. All this reduces the efficiency of the economy.

Deflation, on the other hand, is *inexpedient* because it reduces national income, output and employment. While inflation takes away half the bread of the poor, deflation impoverishes them by taking away the whole of it.

Deflation leads to mass unemployment because fall in production, prices and profits force producers and businessmen to close down their enterprises.

Deflation is also inexpedient because falling prices lead to depression. All economic activities are stagnant. Factories are locked out. Trade and business are at a standstill. There is glut of commodities in all types of markets for goods and services. Even a bumper agricultural crop brings poverty to the peasantry. It is a situation of poverty in the midst of plenty.

Again, once the downward movement of prices begins, the economy plunges into a depression. But the downward movement of the economy is much faster as compared to the upward movement in a cycle. This makes depression of a much longer period. Consequently, people suffer a lot and the economy also remains in a state of stagnation for long.

It is on these grounds that inflation is unjust and deflation is inexpedient.

Keynes pointed out that, “it is not necessary that we weigh one evil against the other. It is easier to agree that both are evils to be shunned.” Still he preferred inflation as the lesser of the two evils. This is because inflation increases national output, employment and income, whereas deflation reduces national income and brings the economy backward to a state of depression. Again inflation is better than deflation because when it occurs the economy is already in a situation of full employment. On the other hand, there is always unemployment under deflation. And unemployment leading to poverty are the two scourges of mankind. Again inflation is a lesser evil than deflation. It redistibutes income and wealth in favour of the rich. But deflation is a greater evil. Though it redistributes income in favour of the low income groups, yet it fails to benefit them because they are unemployed and have little income during deflation. In fact, they are reduced to paupers. It is also easier to control inflation than deflation through appropriate monetary, fiscal, and direct control measures. But to control deflation is a very difficult thing because of the presence of pessimism among producers and businessmen. So far as the increase in inequalities of income and wealth under inflation is concerned, it can be reduced by larger expenditure on social services by the government. The government is in a better position to improve the conditions of the masses under inflation than under deflation due to its larger spending capacity. Moreover, so long as inflation is mild, it helps the economy to grow. It is only when inflation takes the shape of hyperinflation that it is dangerous. Still its effects on the economy may not be so injurious as under deflation.

CONTROL OF DEFLATION

Deflation can be controlled by adopting monetary and fiscal measures in just the opposite manner to control inflation. However, we discuss these measures in brief.

Monetary Policy

To control deflation, the central bank can increase the reserves of commercial banks through a cheap money policy. They can do so by buying securities and reducing the interest rate. As a result, their ability to

extend credit facilities to borrowers increases. But the experience of the Great Depression tells us that in a serious depression when there is pessimism among businessmen, the success of such a policy is practically nil. In such a situation, banks are helpless in bringing about a revival. Since business activity is almost at a stand still, businessmen do not have any inclination to borrow to build up inventories even when the rate of interest is very low. Rather, they want to reduce their inventories by repaying loans already drawn from the banks. Moreover, the question of borrowing for long-term capital needs does not arise during deflation when the business activity is already at a very low level. The same is the case with consumers who faced with unemployment and reduced incomes do not like to purchase any durable goods through bank loans. Thus all that the banks can do is to make credit available but they cannot force businessmen and consumers to accept it. In the 1930s, very low interest rates and the piling up of unused reserves with the banks did not have any significant impact on the depressed economies of the world. Thus the success of monetary policy in controlling deflation is severely limited.

Fiscal Policy

Fiscal policy through increase in public expenditure and reduction in taxes tends to raise national income, employment, output, and prices. An increase in public expenditure during deflation increases the aggregate demand for goods and services and leads to a large increase in income via the multiplier process, while a reduction in taxes has the effect of raising disposable income thereby increasing consumption and investment expenditures of the poeple. The government should increase its expenditure through deficit budgeting and reduction in taxes. The public expenditure includes expenditure on such public works as roads, canals, dams, parks, schools, hospitals and other buildings, etc. and on such relief measures as unemployment insurance, pensions, etc. Expenditure on public works creates demand for the products of private construction industries and helps in reviving them while expenditure on relief measures stimulates the demand for consumer goods industries. Reduction in such taxes as corporate profits tax, income tax, and excise taxes tends to leave more income for spending and investment. Borrowing by the government to finance budget deficits utilises idle money lying with banks and

financial institutions for investment purposes. But the effectiveness of public expenditure primarily depends upon the public works programme, its importance in the economic system, the volume and nature of public works and their planning and timing.

EXERCISES

1. Give a critical asessment of any one theory of inflation and give reasons for selecting this particular theory.
2. What is inflationary-gap? Examine the usefulness of this concept of analysing a process of inflation.
3. Distinguish between demand-pull and cost-push inflation. How have these two views on inflation beeen reconciled?
4. “The distinciton between cost-push and demand-pull inflation is unworkable, irrelevant and even meaningless.” Do you agree with this view. Give reasons in support of your answer.
5. Discuss the theory of structural inflation.
6. Is inflation a purely monetary or purely non-monetary or an institutional phenomenon? Write in the context of a developed country.
7. Discuss the theory of the Phillips curve and bring out its apparent policy implications.
8. How does the Phillips curve expalin the trade-off between unemployment and inflation? Discuss its policy implications.
9. Explain the phenomenon of stagflation. Suggest measures to control it.
10. “Inflation is unjust and deflation is inexpedient.” Discuss.
11. What is inflation? Explain its economic effects on different people.
12. Discuss the causes of inflation.How can it be controlled?
13. Explain Bent Hansen’s Excess Demand Inflation Theory.
14. Explain Friedman's accelerationist hypothesis.
15. Discuss Friedman's view about the Phillips curve.
16. Explain the views of Tobin and Solow with regard to the Phillips curve.
17. Write notes on : Open and Suppressed Inflation, Markup Inflation, Stagflation, Sectoral Inflation, Inflation as a Tax, Costs of Inflation.

CHAPTER 38

BUSINESS CYCLES

MEANING

Business cycle or trade cycle is a part of the capitalist system. It refers to the phenomenon of cyclical booms and depressions. In a business cycle, there are wave-like fluctuations in aggregate employment, income, output and price level. The term business cycle has been defined in various ways by different economists. Prof. Haberler's definition is very simple: "The business cycle in the general sense may be defined as an alternation of periods of prosperity and depression of good and bad trade." Keynes' definition in his *Treatise of Money* is more explicit : "A trade cycle is composed of periods of good trade characterised by rising prices and low unemployment percentage, altering with periods of bad trade characterised by falling prices and high unemployment percentages." Gordon's definition is precise: "Business cycles consist of recurring alternation of expansion and contraction in aggregate economic activity, the alternating movements in each direction being self-reinforcing and prevading virtually, all parts of the economy." The most acceptable definition is by Estey: "Cyclical fluctuations are characterised by alternating waves of expansion and contraction. They do not have a fixed rhythm, but they are cycles in that the phases of contraction and expansion recur frequently and in fairly similar patterns."[1]

TYPES OF BUSINESS CYCLES

Business cycles are usually classified as under :

(1) The Short Kitchin Cycle. It is also known as the minor cycle which is of approximately 40 months duration. It is famous after the name of the British economist Joseph Kitchin, who made a distinction between a major and a minor cycle in 1923. He came to the conclusion on the basic of his research that a major cycle is composed of two or three minor cycles of 40 months.

(2) The Long Jugler Cycle. This cycle is also known as the major cycle. It is defined "as the fluctuation of business activity between successive crises." In 1862 Clement Jugler, French economist showed that periods of prosperity, crisis and liquidation followed each other always in the same order. Later economists have come to the conclusion that a Jugler cycle's duration is on the average nine and a half years.

(3) The Very Long Kondratieff Cycle. In 1925, N.D. Kondratieff, the Russian economist, came to the conclusion that there are longer waves of cycles of more than 50 years duration, made of six Jugler cycles. A very long cycles has come to be known as the Kondratieff wave.

(4) Building Cycles. Another type of cycle relates to the construction of buildings which is of fairly regular duration. Its duration is twice that of the major cycles and is on an average of 18 years' duration. Such cycles are associated with the names of Warren and Pearson.

(5) Kuznets Cycle. Simon Kuznets, propounded a new type of cycle, the secular swing of 16-22 years which is so pronounced that it dwarfs the 7 to 11 years cycle into relative insignificance. This has come to be known as the Kuznets Cycle.

CHARACTERISTICS OF BUSINESS CYCLES

Business cycles possess the following characteristics :

1. Cyclical fluctuations are wave-like movements.

2. Fluctuations are recurrent in nature.

3. They are non-periodic or irregular. In other words, the peaks and

troughs do not occur at regular intervals.

4. They occur in such aggregate variables as output, income, employment and prices.

5. These variables move at about the same time in the same direction but at different rates.

6. The durable goods industries experience relatively wide fluctuations in output and employment but relatively small fluctuations in prices. On the other hand, nondurable goods industries experience relatively wide fluctuations in prices but relatively small fluctuations in output and employment.

7. Business cycles are not seasonal fluctuations such as upswings in retail trade during Diwali or Christmas.

8. They are not secular trends such as long-run growth or decline in economic activity.

9. Upswings and downswings are cumulative in their effects.

Thus business cycles are recurrent fluctuations in aggregate employment, income, output and price level.

---

1. J.A. Estey, *Business Cycles,* 1956.

PHASES OF A BUSINESS CYCLE

A typical cycle is generally divided into four phases : (1) expansion or prosperity or the upswing; (2) recession or upper-turning point; (3) contraction or depression or downswing; and (4) revival or recovery or lower-turning point. These phases are recurrent and uniform in the case of different cycles. But no phase has definite periodicity or time interval. As pointed out by Pigou, cycles may not be twins but they are of the same family. Like families they have *common* characteristics that are capable of description. Starting at the trough or low point, a cycle passes through a

recovery and prosperity phase, rises to a peak, declines through a recession and depression phase and reaches a trough. This is shown in Figure 1 where *E* is the equilibrium position. We describe below these characteristics of a business cycle.

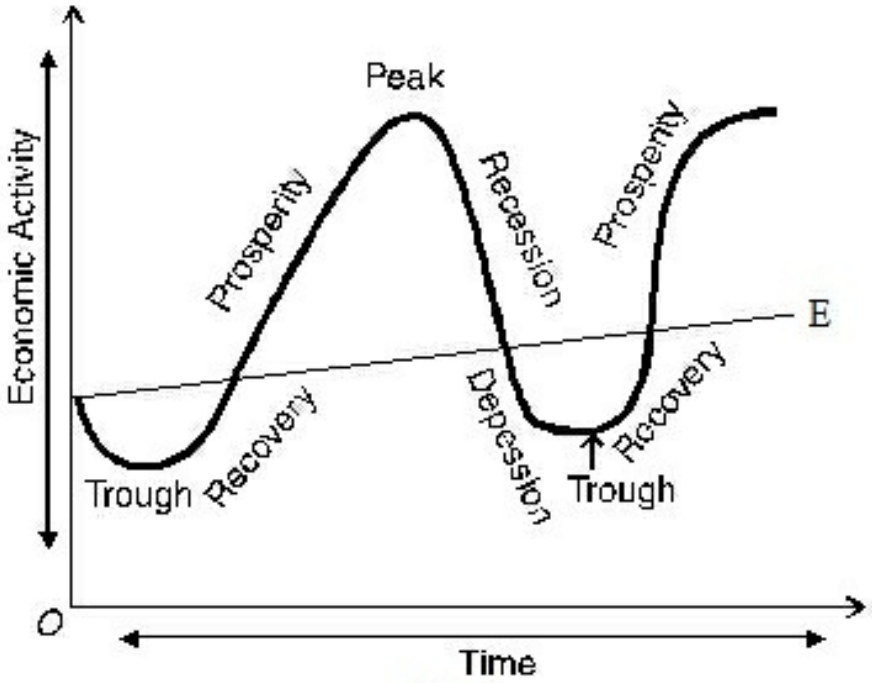


Fig. 1

Recovery

We start from a situation when depression has lasted for some time and the revival phase or the lower-turning point starts. The "originating forces" or "starters" may be exogenous or endogenous forces. Suppose the semi-durable goods wear out which necessitate their replacement in the economy. It leads to increased demand. To meet this increased demand, investment and employment increase. Industry begins to revive. Revival also starts in related capital goods industries. Once begun, the process of revival becomes cumulative. As a result, the levels of employment, income and output rise steadily in the economy. In the early stages of the revival phase, there is considerable *excess* or *idle* capacity in the economy so that output increases without a proportionate increase in total costs. But as time goes on, output becomes less elastic; bottlenecks appear with rising costs, deliveries are more difficult and plants may have to be expanded. Under these conditions, prices rise. Profits increase. Business expectations improve. Optimism prevails. Investment is encouraged which tends to raise the demand for bank loans. It leads to credit expansion. Thus the cumulative process of increase in investment, employment, output, income and prices feeds upon itself and becomes self-reinforcing.

Ultimately, revival enters the prosperity phase.

Prosperity

In the prosperity phase, demand, output, employment and income are at a high level. They tend to raise prices. But wages, salaries, interest rates, rentals and taxes do not rise in proportion to the rise in prices. The gap between prices and costs increases the margin of profit. The increase of profit and the prospect of its continuance commonly cause a rapid rise in stock market values. The economy is engulfed in waves of optimism. Larger profit expectations further increase investment which is helped by liberal bank credit. Such investments are mostly in fixed capital, plant, equipment and machinery. They lead to considerable expansion in economic activity by increasing the demand for consumer goods and further raising the price level. This encourages retailers, wholesalers and manufacturers to add to inventories. In this way, the expansionary process becomes cumulative and self-reinforcing until the economy reaches a very high level of production, known as the *peak* or *boom*.

The peak or prosperity may lead the economy to over full employment and to inflationary rise in prices. It is a symptom of the end of the prosperity phase and the beginning of the recession. The seeds of recession are contained in the boom in the form of strains in the economic structure which act as brakes to the expansionary path. They are : (*i*) scarcities of labour, raw materials, etc. leading to rise in costs relative to prices; (*ii*) rise in the rate of interest due to scarcity of capital; and (*iii*) failure of consumption to rise due to rising prices and stable propensity to consume when incomes increase. The first factor brings a decline in profit margins. The second makes investments costly and along with the first, lowers business expectations. The third factor leads to the piling up of inventories indicating that sales (or consumption) lag behind production. These forces become cumulative and self-reinforcing. Entrepreneurs, businessmen and traders become over cautious and over optimism gives way to pessimism. This is the beginning of the upper turning point.

Recession

Recession starts when there is a downward descend from the ‘peak’ which

is of a short duration. It marks the turning period during which the forces that make for contraction finally win over the forces of expansion. Its outward signs are liquidation in the stock market, strain in the banking system and some liquidation of bank loans, and the beginning of the decline of prices. As a result, profit margins decline further because costs start overtaking prices. Some firms close down. Others reduce production and try to sell out accumulated stocks. Investment, employment, income and demand decline. This process becomes cumulative.

Recession may be mild or severe. The latter might lead to a sudden explosive situation emanating from the banking system or the stock exchange, and a *panic* or *crisis* occurs. “When a crisis, and more particularly a panic, does occur, it seems to be associated with a collapse of confidence and sudden demands for liquidity. This crisis of nerves may itself be occasioned by some spectacular and unexpected failure. A firm or a bank, or a corporation announces its inability to meet its debts. This announcement weakens other firms and banks at a time when ominous signs of distress are appearing in the economic structure; moreover, it sets off a wave of fright that culminates in a general run on financial institutions"...Such was the experience of the United States in 1873, in 1893, in 1907 and recently in 2008. In the words of M.W. Lee, “A recession, once started, tends to build upon itself much as forest fire, once under way, tends to create its own draft and give internal impetus to its destructive ability.”[2]

Depression

Recession merges into depression when there is a general decline in economic activity. There is considerable reduction in the production of goods and services, employment, income, demand and prices. The general decline in economic activity leads to a fall in bank deposits. Credit expansion stops because the business community is not willing to borrow. Bank rate falls considerably. According to Professor Estey, “This fall in active purchasing power is the fundamental background of the fall in prices, that despite the general reduction of output, characterises the depression.” Thus a depression is characterised by mass unemployment; general fall in prices, profits, wages, interest rate, consumption,

expenditure, investment, bank deposits and loans; factories close down; and construction of all types of capital goods, buildings, etc. comes to a standstill. These forces are cumulative and self-reinforcing and the economy is at the *trough*.

The trough or depression may be short-lived or it may continue at the bottom for considerable time. But sooner or later limiting forces are set in motion which ultimately tend to bring the contraction phase to end and pave the way for the revival. A cycle is thus complete.

CAUSES OF BUSINESS CYCLES

During the last several hundred years, philosophers, economists, stock brokers and men in the street have tried to give various causes of business cycles. Some attribute them to monetary and non-monetary factors while others to psychological factors. Samuelson attributes business cycles to external and internal factors which we explain below.

External Factors

The external factors emphasise the causes of business cycles in the fluctuations of something outside the economic system. Such external factors are sunspots, wars, revolutions, political events, gold discoveries, growth rate of population, migrations, discoveries and innovations. These outside factors change the level of national income by affecting either the investment or consumption component of aggregate demand. For example, a drought that destroys many crops due to sunspots may reduce the quantity of goods produced in the country and adversely affect both consumption and investment. An innovation by opening the door to new markets, raw materials, products and production processes encourages new investments in plant and equipment. Inventions of railroads, electricity, telephone, automobiles, TVs, computers, etc. have led to the burst of investments in both capital and consumer goods from time to time. Discoveries of gold, oil and natural resources have led to large scale investments. Similarly, population expansion and migrations are the causes of huge investments in both housing and other infrastructure and consumer durables. All the above noted external factors have been

responsible for booms in business cycles from time to time.

Internal Factors

The internal factors relate to "mechanisms within the economic system itself which will give rise to self-generating business cycles, so that every expansion will breed recession and contraction, and every contraction will in turn breed revival and expansion, in a quasi-regular, repeating, never-ending chain."[3] Haberler divides the internal factors into monetary and non-monetary which we briefly explain.

1. Bank Credit. Hawtrey, Friedman and other monetarists regard business cycles as "a purely monetary phenomenon". According to Hawtrey, cyclical fluctuations are caused by expansion and contraction of bank credit. These in, turn, lead to changes in the demand for money on the part of producers and traders. Bank credit is the principal means of payment. Credit is expanded or reduced by the banks by lowering or raising the rate of interest or by purchasing and selling of securities to traders. This increases or decreases the supply of money in the economy. An increase in the money supply brings about prosperity and a decrease in the money supply leads to depression.

---

2. M.W. Lee, *Economic Fluctuations*, 1955.

2. Over-Saving or Under-Consumption. According to economists like Hobson, Foster and Douglas, business cycles are caused by oversaving or under-consumption. They argue that wide disparities of income and wealth lead to depression in the country. The rich people are not able to spend their entire income. So they save more and invest more in producing consumer goods. On the other hand, the poor people have low incomes or wages. As a result, their demand for consumer goods is low which means that there is under-consumption. According to Hobson, oversaving leads to production of consumer goods in large quantities and to a boom. But under-consumption on the part of the workers due to low wages brings a fall in the demand for consumer goods. Stocks pile up at the current level of prices. These, in turn, lead to a fall in the prices of consumer goods and in the income of the producers. As a result,

depression sets in.

3. Over-Investment. Hayek, Spiethoff, Cassel and Robertson find the root cause of business cycles in over-investment. According to Hayek, it is bank loans which lead to over-investment in capital goods industries relative to consumer goods industries that ultimately brings depression in the economy. When the total money supply exceeds the amount of voluntary savings, it leads to increase in the investment activity and ultimately to a boom. But banks cannot continue to give credit for long due to the shortage of voluntary savings. As a result, production will decline which will bring about a depression. Thus it is over-investment in the capital goods industries which is the cause of a boom and a depression.

4. Competition. According to Chapman, the main cause of business cycles is the existence of competition in an economy which leads to over-production and ultimately to a crisis (depression). Under competitive conditions, firms produce in anticipation of demand. The profit motive attracts new firms. Production increases and boom starts. Competition and profits lead to overproduction and glut of commodities in the market and to fall in prices. On the other hand, the race to produce more and profit more on the part of producers increases the demand for factors of production. Competition among producers to hire more factors raises their prices. Thus costs rise which raise the prices of products. Demand falls and there is glut of commodities which eventually leads to fall in prices and to a depression.

5. Psychological Causes. According to Pigou, the alternating waves of "over optimism" and "over pessimism" are the sole causes of the industrial fluctuations. He traces cyclical fluctuations to the tendency of businessmen to react excessively to the changing conditions of the economy. It is this tendency that causes alternating periods of over-production and under-production. Errors of optimism and pessimism are interacting forces. As soon as the business community discovers that it has made an error of optimism, it tries to correct it by making errors of pessimism. Each phase of the cycle produces a state of psychology which produces forces that bring about reversal of that psychology and in turn

another reversal. These alternating waves of over-optimism (over-production) and over-pessimism (under-production), as a result of these reversals, are the main causes of business cycles.

---

3. P.A. Samuelson, *Economics*, p. 293

6. Innovations. According to Schumpeter, innovations in the structure of an economy are the source of economic fluctuations. To him, "the cause of depression is prosperity." The boom consists in the carrying out of innovations in the industrial and commercial fields. The cyclical upswing is set in motion when an innovator starts making investment in his innovation of a new product. This enables him to make profit. Soon other enrepreneurs adopt this new product in "swarm-like clusters". Innovations in one field induce innovations in related fields. There is large increase in the output of new products. Consequently, money incomes and prices rise and help to create a cumulative expansion in the economy. Overoptimism adds further to the boom. When there is glut of new products in the market, their prices fall, and profit margins of entrepreneurs are reduced. Banks ask for repayment of loans. The quantity of money is reduced and prices fall further. Some entrepreneurs cut down production and others are forced into liquidation. Thus the economy enters into depression.

7. Marginal Efficiency of Capital (MEC). According to Keynes, the cycle consists primarily of fluctuations in the rate of investment. And fluctuations in the rate of investment are caused mainly by fluctuations in the MEC. The MEC depends on the supply price of capital assets and their prospective yield. The supply price of capital assets being stable in the short-run, the MEC is determined by the prospective yield of capital assets. The prospective yield, in turn, depends on business expectations. Fluctuations in the rate of investment are also caused by fluctuations in the rate of interest. But it is fluctuations in the MEC which are the principal cause of cyclical fluctuations.

Conclusion. To conclude with Samuelson, business cycles are caused both by external and internal factors. The economic system responds to fluctuations in external factors according to its internal factors, and vice versa.

EFFECTS OF BUSINESS CYCLES

Business cycles have both good and bad effects depending upon whether the economy is passing through a phase of prosperity or depression.

In the *prosperity* phase, "the real income consumed, real income produced and the level of employment are high or rising and there are no idle or unemployed workers or very few of either." There is general increase in economic activity : aggregate output, demand, employment and income are at a high level. Prices are rising. Profits are increasing. Stock markets are rapidly reaching new heights. Investments are increasing with liberal bank credit. This entire process is cumulative and self-reinforcing.

But different sections of the society are affected differently during the prosperity phase. The landless, factory and agricultural workers and middle classes suffer because their wages and salaries are more or less fixed but the prices of commodities rise continuously. They become more poor. On the other hand, businessmen, traders, industrialists, real estate holders, speculators, landlords, shareholders and others with variable incomes gain. Thus the rich become richer and the poor poorer.

The social effects are also bad. Lured by profit, there is hoarding, black-marketing, adulteration, production of substandard goods, speculation, etc. Corruption spreads in every walk of life.

When the economy is nearing the *full employment level* of resources, the ill-effects on production start appearing. Rising prices of raw materials and increase in wages raise costs of production. As a result, profit margins decline. There is rise in interest rates due to scarcity of capital which makes investment costly. These two factors lower business expectations. Lastly, the demand for consumer goods does not rise due to inflationary rise in prices. This leads to piling up of inventories (stocks) with producers and traders. Thus sales lag behind production. There is decline is prices. Producers, businessmen and traders become pessimists and the recession starts.

During *recession*, profit margins decline further because costs start rising

more than prices. Some firms close down. Others reduce production and try to sell accumulated stocks. Investment, output, employment, income, demand and prices decline further. This process becomes cumulative and recession merges into depression.

During a *depression*, there is mass unemployment. Prices, profits and wages are at their lowest levels. Demand for goods and services is the minimum. Investment, bank deposits and bank loans are negligible. Construction of all types of capital goods, buidlings, etc. is at a standstill. There is mass unemployment in the economy. The government revenues from direct and indirect taxes decline. The real burden of the debt increases. The economic development of the country suffers.

THEORIES OF BUSINESS CYCLES

In fact, the causes of business cycles given above are based on the theories of business cycles propounded by economists from time to time. We discuss some of the important theories as under.

1. Hawtrey's Monetary Theory

According to Prof. R.G. Hawtrey,[4] "The trade cycle is a purely monetary phenomenon." It is changes in the flow of monetary demand on the part of businessmen that lead to prosperity and depression in the economy. He opines that non-monetary factors like strikes, floods, earthquakes, droughts, wars, etc. may at best cause a partial depression, but not a general depression. In actuality, cyclical fluctuations are caused by expansion and contraction of bank credit which, in turn, lead to variations in the flow of monetary demand on the part of producers and traders. Bank credit is the principal means of payment in the present times. Credit is expanded or reduced by the banking system by lowering or raising the rate of interest or by purchasing or selling securities to merchants. This increases or decreases the flow of money in the economy and thus brings about prosperity or depression.

The *expansion phase* of the trade cycle starts when banks increase credit facilities. They are provided by reducing the lending rate of interest and

by purchasing securities. These encourage borrowings on the part of merchants and producers. This is because they are very sensitive to changes in the rate of interest. So when credit becomes cheap, they borrow from banks in order to increase their stocks or inventories. For this, they place larger orders with producers who, in turn, employ more factors of production to meet the increasing demand. Consequently, money incomes of the owners of factors of production increase, thereby increasing expenditure on goods. The merchants find their stocks being exhausted. They place more orders with producers. This leads to further increase in productive activity, income, outlay, and demand, and a further depletion of stocks of merchants. According to Hawtrey, "Increased activity means increased demand, and increased demand means increased activity. A vicious circle is set up, a cumulative expansion of productive activity."

---

4. R.G. Hawtrey, *Trade and Credit,* 1928

As the cumulative process of expansion continues, producers quote higher and higher prices. Higher prices induce traders to borrow more in order to hold still larger stocks of goods so as to earn more profits. Thus optimism encourages borrowing, borrowing increases sales, and sales raise optimism.

According to Hawtrey, prosperity cannot continue limitlessly. It comes to an end when banks stop credit expansion. Banks refuse to lend further because their cash funds are depleted and the money in circulation is absorbed in the form of cash holdings by consumers. Another factor is the export of gold to other countries when imports exceed exports as a result of high prices of domestic goods. These factors force the banks to raise interest rates and refuse to lend. Rather, they ask the business community to repay their loans. This starts the recessionary phase.

In order to repay bank loans, businessmen start selling their stocks. This sets the process of falling prices. They also cancel orders with producers. The latter curtail their productive activities due to fall in demand. These, in turn, lead to reduction in the demand for factors of production. There is unemployment. Incomes fall. Falling demand, prices and incomes are the

signals for depression. Unable to repay bank loans, some firms go into liquidation, thus forcing banks to contract credit further. Thus the entire process becomes cumulative and the economy is forced into depression.

According to Hawtrey, the *process of recovery* is very slow and halting. As *depression* continues, traders repay bank loans by selling their stocks at whatever prices they can. As a result, money flows into the reserves of banks and funds increase with banks. Even though the bank rate is very low, there is "credit deadlock" which prevents businessmen to borrow from banks due to pessimism in economic activity. This deadlock can be broken by following a cheap money policy by the central bank which will ultimately bring about recovery in the economy.

Its Criticisms

Monetarists like Friedman have supported Hawtrey's theory. But the majority of economists have criticised him for over-emphasising monetary factors to the neglect of non-monetary factors in explaining cyclical fluctuations. Some of the points of criticism are discussed below.

(1) Credit not the Cause of Cycle. None can deny that expansion of credit leads to the expansion of business activity. But Hawtrey believes that an expansion of credit leads to a boom. This is not correct because the former is not the cause of the latter. As pointed out by Pigou, "Variations in the bank money supply is a part of the business cycle, it is not the cause of it." At the bottom of the depression, credit is easily available. Even then, it fails to bring a revival. Similarly, contraction of credit cannot bring about a depression. At best, it can create conditions for that. Thus expansion or contraction of credit cannot originate either boom or depression in the economy.

(2) Money Supply cannot Continue a Boom or Delay a Depression. Haberler has criticised Hawtrey for "his contention that the reason for the breakdown of the boom is always a monetary one and that prosperity could be prolonged and depression stayed off indefinitely if the money suply were inexhaustible."[5] But the fact is that even if the supply of money is inexhaustible in the country, neither prosperity can be continued indefinitely nor depression can be delayed idefinitely.

(3) Traders do not Depend Only on Bank Credit. Hamberg[6] has criticised Hawtrey for the role assigned to wholesalers in his analysis. The kingpin in Hawtrey's theory is the trader or the wholesaler who gets credit from banks and starts the upturn or vice-versa. In actuality, traders do not depend exclusively on bank credit but they finance business through their own accumulated funds and borrowing from private sources.

(4) Traders do not React to changes in Interest Rates. Further, Hamberg also does not agree with Hawtrey that traders react to changes in interest rates. According to Hamberg, traders are likely to react favourably to a reduction in the interest rate only if they think that the reduction is permanent. But they do not react favourably during the depression phase because traders expect a further reduction every time the interest rate is reduced. On the other hand, if traders finance their stocks with their own funds, interest rate changes will have little effect on their purchases.

(5) Factors other than Interest Rate More Important. It is an exaggeration to say that the decisions of traders regarding accumulation or depletion of stocks are solely governed by changes in interest rate. As a matter of fact, factors other than the rate of interest are more important in influencing such decisions. They are business expectations, price changes, cost of storage, etc.

(6) Inventory Investments do not Produce True Cycles. Hamberg further points out that in Hawtrey's theory cumulative movements in economic activity are the result of changes in stocks of goods. But fluctuations in inventroy investment can at best produce minor cycles which are not cycles in the true sense of the term.

(7) Does not Explain Periodicity of Cycle. The theory also fails to explain the periodicity of the cyle.

(8) Ignores Non-Monetary Factors. Hawtrey's theory is incomplete because it emphasises only monetary factors and totally ignores such non-monetary factors as innovations, capital stock, multiplier-accelerator interaction, etc.

2. Hayek's Monetary Over-Investment Theory

F.A. Hayek formulated his monetary over-investment theory of trade cycle. He explained his theory on the basis of Wicksell's distinction between the natural interest rate and the market interest rate. The natural rate of interest is that rate at which the demand for loanable funds equals the supply of voluntary savings. On the other hand, the market rate of interest is the money rate which prevails in the market and is determined by the demand and supply of money. According to Hayek, so long as the natural rate of interest equals the market rate of interest, the economy remains in the state of equilibrium and full employment . Trade cycles in the economy are caused by inequality between market and natural interest rates. When the market interest rate is less than the nautral rate, there is prosperity in the economy. On the contrary, when the market interest rate is more than the natural rate, the economy is in depression.

---

5. G. Haberler, *Business Cycles*, pp. 27-28.

6. D. Hamberg, *Business Cycles*, 1937.

According to this thoery, prosperity begins when the market rate of interest is less than the natural rate of interest. In such a situation, the demand for investment funds is more than the supply of available savings. The demand for investment funds is met by the increase in the supply of money. As a result, the interest rate falls. Low interest rate induces producers to get more loans from banks. The producers get more loans to invest for the production of more capital goods. They adopt capital-intensive methods for producing more of capital goods. As a result, production costs fall and profits increase. The production process becomes very lengthy with the adoption of capital-intensive methods. This has the effect of increasing the prices of capital goods in comparison to consumer goods. There being full employment in the economy, they transfer factors of the production from consumer goods sector to capital goods sector. Consequently, the production of consumer goods falls, their prices increase and their consumption decreases. Forced savings increase with the fall in consumption which are invested for the production of capital goods. This leads to increase in their production. On the other hand, with increase in the prices of consumer goods, their producers earn more profits. Induced by high profits, they try to produce more. For this, they

pay higher remuneration to factors of production in comparison with the producers of capital goods. There being competition between the two sectors, prices of factors and prices in the economy continue to rise. This leads to the atmosphere of prosperity in the country and monetary over-investment on factors spreads the boom.

According to Hayek, when the prices of factors are rising continously, the rise in production costs bring fall in profits of producers. The producers of capital goods invest less in the expectation of loss in the future. Consequently, the natural interest rate falls. Simultaneously, banks impose restrictions on giving loans to them. With low profits and reduction in loans, producers reduce the production of capital goods and adopt labour-intensive production processes. There is less investment in capital goods. Production process being small and labour-intensive, the demand for money is reduced, which increases the market interest rate which is more than the natural interest rate. Producers transfer the factors from the production of capital goods to that of consumer goods. But more factors cannot be used in the consumer goods sector as compared to the capital goods sector. This leads to fall in the prices of factors and resources become unemployed. Thus, with the continuous reduction in the prices of goods and factors in the economy, a long period of depression and unemployment begins.

According to Hayek, when the fall in prices comes to an end during depression, banks begin to raise the supply of money which reduces the market interest rate below the natural interest rate. This encourages investment and the process of revival begins in the economy.

Its Criticisms

The monetary over-investment theory of Hayek has been criticised on the following counts:

(1) Narrow Assumption of Full Employment. This theory is based on the assumption of full employment according to which capital goods are produced by reducing consumer goods. In reality, there is no full employment of resources. If resources remain unutilised, the expansion of both the capital goods sector and consumer goods sector may occur

simultaneously. In such a situation, there is no need of transferring resources from one sector to the other.

(2) Unrealistic Assumption of Equilibrium. The assumption of this theory that in the beginning savings and investment are in equilibrium in the economy and the banking system destroys this equilibrium is unrealistic. This is because the equilibrium may deviate due to both internal and external reasons.

(3) Interest Rate not the only Determinant. Hayek assumes changes in the rate of interest as the cause of fluctuations in the economy. This is not correct because besides changes in the rate of interest, the expectations of profit, innovation, invention, etc. also affect trade cycles.

(4) Undue Importance to Forced Savings. Prof. Strigl has criticised this theory for giving undue importance to forced savings. According to him, when people with fixed incomes reduce their consumption with the increase in prices and the high income groups also reduce their consumption to the same extent, savings will not be forced but voluntary.

(5) Investment does not fall with Increase in Consumer Goods. Hayek argues that with the production of consumer goods and the increase in profits from them, investment falls in capital goods. This is not correct. According to Keynes, the marginal productivity of capital increases with the increase in profits of consumer goods. As a result, investment in capital goods also increases and does not fall.

6. Incomplete Theory. Hayek's theory is incomplete because it does not explain the various phases of trade cycle.

3. Schumpeter's Innovations Theory

The innovations theory of trade cycles is associated with the name of Joseph Schumpeter.[7] According to Schumpeter, innovations in the structure of an economy are the source of economic fluctuations. Trade cycles are the outcome of economic development in a capitalist society. Schumpeter accepts Juglar's statement that "the cause of depression is prosperity," and then gives his own view about the originating cause of

the cycle.

Schumpeter’s approach involves the development of his model into two stages. The first stage deals with the initial impact of innovation and the second stage follows through reactions to the original impact of innovation.

The first approximation starts with the economic system in equilibrium with every factor fully employed. Every firm is in equilibrium and producing efficiently with its costs equal to its receipts. Product prices are equal to both average and marginal costs. Profits and interest rates are zero. There are no savings and investments. This equilibrium is characterised by Schumpeter as the "circular flow" which continues to repeat itself in the same manner year after year, similar to the circulation of the blood in an animal organism. In the circular flow, the same products are produced every year in the same manner.

Schumpeter’s theory starts with the breaking up of the circular flow by an innovation in the form of a new product by an entrepreneur for earning profit.

By innovation Schumpeter means “such changes in the production of goods as cannot be affected by infinitesimal steps or variations on the margin.” An innovation may consist of : (1) the introduction of a new product; (2) the introduction of a new method of production; (3) the opening up of a new market; (4) the conquest of a new source of raw materials or semi-manufactured goods; and (5) the carrying out of the new organisations of an industry. Innovations are not inventions. According to Schumpeter, there is nothing that can explain that inventions occur in a cyclical manner. It is the introduction of a new product and the continual improvements in the existing ones that are the principal causes of business cycles.

Schumpeter assigns the role of an innovator not to the capitalist but to an entrepreneur. The entrepreneur is not a man of ordinary ability but one who introduces something entirely new. He does not provide funds but directs their use.

To perform his economic function, the entrepreneur requires two things: first, the existence of technical knowledge in order to produce new products, and second, the power of disposal over the factors of production in the form of bank credit. According to Schumpeter, a reservoir of untapped technical knowledge exists in a capitalist society which he can make use of. Therefore, credit is essential for breaking the circular flow.

The innovating entrepreneur is financed by expansion of bank credit. Since investment in an innovation is risky, he must pay interest on it. With his newly acquired funds, the innovator starts bidding away resources from other industries. Money incomes increase. Prices begin to rise, thereby stimulating further investment. The new innovation starts producing goods and there is an increased flow of goods in the economy. Consequently, supply exceeds demand. Prices and cost of production of goods start declining until recession sets in. Because of the low prices of goods, producers are not willing to expand production. During this period of recession, credit, prices and interest rate decline but total output is likely to average larger than in the preceding prosperity.

Thus Schumpeter's first approximation consists of a two-phase cycle. The economy starts at the equilibrium state, rises to a peak and then starts downward into a recession and continues till the new equilibrium is reached. This new equilibrium will be at a higher level of income than the initial equilibrium because of the innovation which started the cycle. This is shown as the “Primary Wave” in Figure 2.

The second approximation of Schumpeter follows through the reaction of the impact of original innovation. Once the original innovation becomes successful and profitable, other entrepreneurs follow it in "swarm-like clusters." Innovation in one field induces innovations in related fields. Consequently, money incomes and prices rise and help to create a cumulative expansion throughout the economy. With the increase in the purchasing power of consumers, the demand for the products of old industries increases in relation to supply. Prices rise further. Profits increase and old industries expand by borrowing from the banks. It induces a *secondary* wave of credit inflation which is superimposed on the primary wave of innovation. Overoptimism and speculation add further to

the boom. After a period of gestation, the new products start appearing in the market displacing the old products and enforcing a process of liquidation, readjustment and absorption.

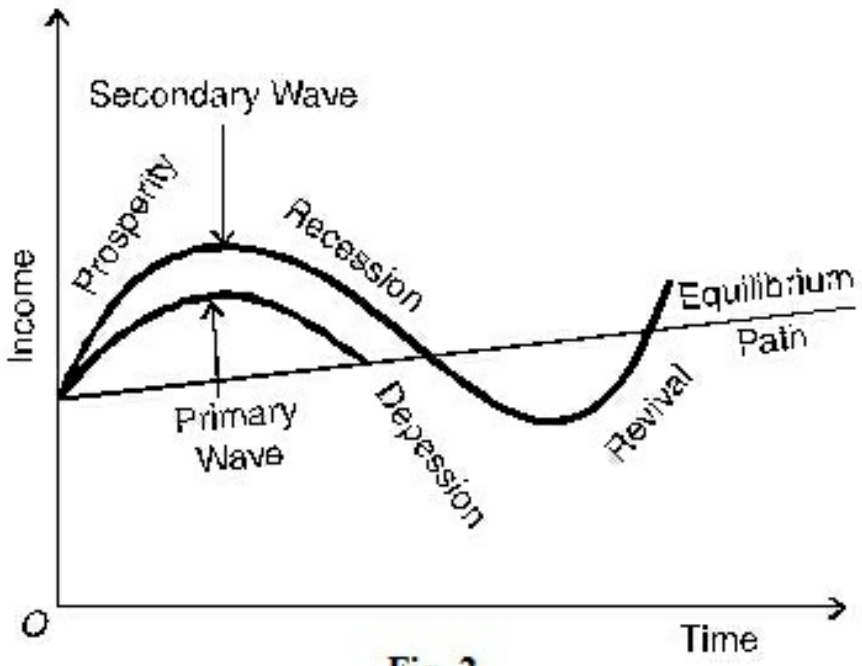


Fig. 2

The demand for the old products is decreased. Their prices fall. The old firms contract output and some are even forced to run into liquidation. As the innovators start repaying bank loans out of profits, the quantity of money is decreased and prices tend to fall. Profits decline. Uncertainty and risks increase. The impulse for innovation is reduced and eventually comes to an end. Depression sets in, and the painful process of readjustment to the "point of previous neighbourhood of equilibrium" begins. Ultimately, the natural forces of recovery bring about a revival.

---

7. J.A. Schumpeter, *Business Cycles*, 1939.

Schumpeter believes in the existence of Kondratieff long wave of upswings and downswings in economic activity. Each long wave upswing is brought about by an innovation which leads to abundance of goods for the masses. Once the upswing ends, the long wave downswing begins.

Thus the second approximation of Schumpeter's theory of trade cycle develops into a four phase cycle with the recession which was the second phase in the first approximation continuing downward to give the depression phase. This extension of cycle is followed by a period of revival which continues till the equilibrium level is reached. This is shown as the "Secondary Wave" in Figure 2.

Its Criticisms

Schumpeter’s treatment of the different phases and turning points of the cycle is novel and different from all other economists. But it is not free from certain criticisms.

(1) Innovator not Necessary for Innovations. Schumpeter's analysis is based on the innovator. Such persons were to be found in the 18th and 19th centuries who made innovations. But now all innovations form part of the functions of joint stock companies. Innovations are regarded as the routine of industrial concerns and do not require an innovator as such.

(2) Innovations not the Only Cause of Cycles. Schumpeter’s contention that cyclical fluctuations are due to innovations is not correct. As a matter of fact, trade cycles may be due to psychological, natural or financial causes.

(3) Bank Credit not the Only Source of Funds. Schumpeter gives too much importance to bank credit in his theory. Bank credit may be important in the short run when industrial concerns get credit facilities from banks. But in the long run when the need for capital funds is much greater, bank credit is insufficient. For this, business houses have to float fresh shares and debentures in the capital market. Schumpeter's theory is weak in that it does not take these factors into consideration.

(4) Innovation financed through Voluntary Savings does not produce a Cycle. Critics point out that if an innovation is financed through voluntary savings or internal funds, there will not be an inflationary rise in prices. Consequently, in an underemployed economy an innovation financed through voluntary savings might not generate a cycle.

(5) Full Employment Assumption Unrealistic. Schumpeter's analysis is based on the unrealistic assumption of full employment of resources to begin with. But the fact is that at the time of revival, the resources are unemployed. Thus the introduction of an innovation may not lead to the withdrawal of labour and other resources from old industries. Thus the competitive impact of an innovation would not increase costs and prices. Since full employment is an exception rather than the rule. Thus

Schumpeter's theory is not a correct explanation of trade cycles.

4. The Psychological Theory

The psychological theory of business cycle has been mainly developed by Prof. A.C. Pigou.[8] This theory attempts to explain the phenomenon of business cycle on the basis of changes in the psychology of industrialists and businessmen. The tendency of the business class is to react excessively to the changing conditions of the economy that are mainly responsible for cyclical fluctuations.

According to Pigou, expectations originate from some real factors such as good harvests, wars, natural calamities, industrial disputes, innovations, etc. But he attributes the causes of business cycle into two categories : (a) inpulses and (b) conditions. *Impulses* refer to those causes which set a process in motion. The *conditions,* on the other hand, are the vehicles through which the process passes and upon which the impulses act. These conditions are the decision making centres which, in turn, shift the levels of economic activities and bring necessary changes in their compositions. They include monetary institutions, market structures, trade unions, etc.

Pigou divides impulses into two parts : (*i*) The expectations held by businessmen, and (*ii*) the actual economic resources owned by them. The expectations depend upon the psychology of businessmen and on their control over resources. But expectations which correspond to actual changes in the economy and are realised, they do not generate cyclical fluctuations.

According to Pigou, it is only when expectations are devoid of their realistic basis, there may be error in forecasting. Such type of expectations cause disturbances in the economy and result in waves of optimism and pessimism. Such “errors in forecasting” may be due to : (*i*) the deviation of actual demand from anticipated demand on the part of consumers; (*ii*) the continual and unpredictable change in the values of economic variables, and (iii) the existence of time lags on account of gestation periods.

Once an error of forecasting occurs in any sector of the economy, it

spreads in the same directions. Once this “impulse” starts acting on the “conditions”, it feeds upon itself. According to Pigou, this is because there is a certain measure of psychological interdependence. In other words, the expectations of optimism or pessimism on the part of businessmen strengthen the building up of further expectations of the same type. When businessmen have a feeling of optimism about the future prospects of business, it would increase the demand for investment resources and inter-industry relations would induce businessmen in other industries to be optimist. Consequently, there is the emergence of boom conditions in the economy.

Pigou opines that the wave of optimism is replaced by pessimism on account of time lags in production. Being over optimistic, some producers make the mistake of over investing in goods. When the goods start coming into the market in large quantities, it is not possible to sell them at remunerative prices. As a result, inventories accumulate. A wave of pessimism starts which spreads to other sectors of the economy. This leads to the emergence of slump in the country. To Pigou, *the lower turning point* starts when inventories are depleted and the "bolder spirit of industry" helps to revive expectations. As a result, the rays of optimism spread slowly and revival starts which leads to boom and so on.

---

8. A.C. Pigou, *Industrial Fuctuations*, 1929.

Thus according to this theory, booms and slumps are due to alternative waves of optimism and pessimism on the part of businessmen and industrialists.

Its Criticisms

The psychology theory has been criticised for the following reasons:

1. This is not a theory of business cycles in the true sense because it fails to explain the different phases of a business cycle.

2. It fails to explain the periodicity of a business cycle.

3. It neglects the role of various exogenous and monetary factors which

influence business expectations.

4. The theory does not explain fully the causes that give rise ‘to waves of optimism and pessimism’ in the business world.

5. The theory fails to explain the reason for deficiency of demand when goods start entering the market in larger quantities. Moreover, it does not explain as to why the deficiency of demand overtakes the flow of goods in the market.

5. The Cobweb Theory

The cobweb theory of business cycles was propounded in 1930 independently by Professors H. Schultz of America, J. Tinbergen of the Netherlands and U.Ricci of Italy. But it was Prof. N. Kaldor of Cambridge University, England, who used the name Cobweb Theorem because the pattern of movements of prices and outputs resembled a cobweb.

The cobweb model is used to explain the dynamics of demand, supply and price over long periods of time. There are many perishable agricultural commodities whose prices and outputs are determined over long periods and they show cyclical movements. As prices move up and down in cycles, quantities produced also seem to move up and down in a counter-cyclical manner. Such cycles in commodity prices and outputs are explained in terms of the cobweb model, so called because the diagrams look like cobwebs.

Suppose the production process spreads over two periods: current and previous. Production in the current period is assumed to be determined by decisions made in the previous period. Thus the current output reflects a production decision made by the producer during the previous period. This decision is in response to the price that he *expects* to rule during the current period when the crop is available for sale. But he expects that the price that would be established during the current period would equal the price during the previous period.

The cobweb theory analyses the movements of prices and outputs when supply is wholly determined by prices in the previous period. In order to

find out the conditions for converging, diverging or constant cycles, one has to look first at the slope of the demand curve and then of the supply curve. If the slope of the demand curve is numerically smaller than the slope of the supply curve, the price will converge towards equilibrium. Conversely, if the slope of the demand curve is numerically greater than the slope of supply curve, the price will diverge from equilibrium. If the slope of the demand curve is numerically the same as that of the supply curve, the price will oscillate around its equilibrium value.

Its Assumptions

The cobweb theory is based on the following assumptions:

(1) The current year's (*t*) supply depends upon the last (previous) year's (*t*–1) decisions regarding output level. Hence current output is influenced by last year's price, *i.e*. P (*t*–1).

(2) The current period or year is divided into sub-periods of a week or fortnight.

(3) The parameters determining the supply function have constant values over a series of periods.

(4) Current demand (D_t) for the commodity is a function of current price (P_t).

(5) The price expected to rule in the current period is the actual price in the last year.

(6) The commodity under consideration is perishable and can be stored only for one year.

(7) Both supply and demand functions are linear.

The Theory

There are three types of cobwebs: (1) Convergent; (2) Divergent; and (3) Continuous. They are explained as under:

1. Convergent Cobweb. Under this formulation of the *cobweb theorem*, the supply function is $S_t = S\,(t–1)$ and the demand function is $D_t = D\,(P_t)$. The market equilibrium will be when the quantity supplied equals the quantity demanded: $S_t = D_t$. In any market in which producers' current supply is in response to the price during the last year, equilibrium can be established only through a series of adjustments that take place over several consecutive periods.

Let us take potato growers who produce only one crop a year. They decide about how many potatoes they will grow this year on the assumption that the price of potatoes this year will equal the price in the last year. The market demand and supply curves for potatoes are represented by *D* and *S* curves respectively in Figure 3. The price in the last year was *OP* and the producers decide the equilibrium output *OQ* this year. But the potato crop is damaged due to a blight so that their current ouput is OQ_1 which is smaller than the equilibrium output *OQ*. This leads to rise in the price to OP_1 in the current period. In the next period, the potato growers will produce OQ_2 quantity in response to the higher price OP_1 (=Q_1b). But this is more than the equilibrium quantity *OQ* which is needed in the market. It will, therefore, lower the price to OP_2 (=Q_2d) and thus again lead to changes in the production plans of producers whereby they will reduce supply to OQ_3 in the third period. But this quantity is less than the equilibrium quantity *OQ*. Price will, therefore, rise to OP_3 (=Q_3f) which, in turn, will encourage producers to produce *OQ* quantity. Ultimately, the equilibrium will be established at point *g* where *D* and *S* curves intersect. The series of adjustments just described trace out a cobweb pattern *a, b, c,*

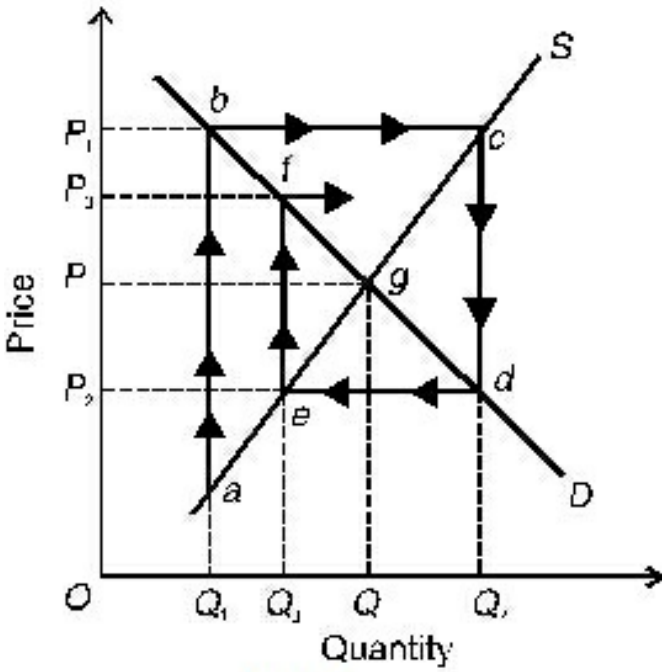


Fig. 3

d, *e* and *f* which converge towards the point of market equilibrium *g* when period-to-period changes in price and quantity have been reduced to zero. The cobweb is convergent.

(2) Divergent Cobweb. But there may be an *unstable* cobweb when price and quantity changes move away from the equilibrium position. This is illustrated in Figure 4. Suppose from the initial price-quantity equilibrium situation of *OP* and *OQ*, there is a temporary disturbance that causes output to fall to OQ_1. This raises the price to OP_1 ($=Q_1a$). The increased price, in turn, raises output to OQ_2 which is more than the equilibrium output *OQ*. Consequently, the price falls to OP_2. But at this price the demand (OQ_2) exceeds the supply (OQ_3). As a result, the price shoots up to OP_3 ($=Q_3e$) and the adjustment of producers to this price leads farther away from the equilibrium. This is an explosive situation and the equilibrium position is unstable. The cobweb is divergent.

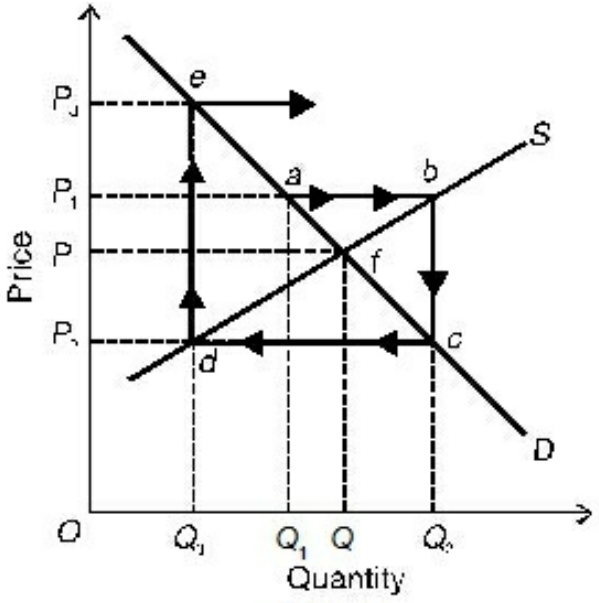


Fig. 4

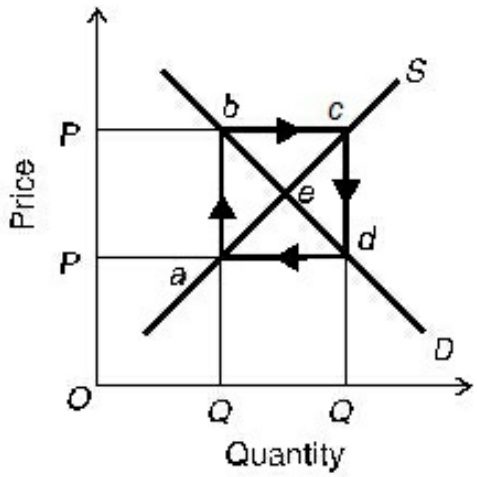


Fig. 5

(3) Continuous Cobweb. The cobweb may be of constant amplitude with perpetually oscillating prices and quantities, as shown in Figure 5. Suppose that the price in the current year is *OP*. Thus the quantity to be supplied is going to be OQ_1. But in order to sell this output, the price that it will fetch in the next period will be OP_1. But at this price, the demand

OQ_1 is more than the supply OQ which will again raise the price to OP ($=Qb$). In this way, prices and quantities will move in a circle with oscillations of constant amplitude around the equilibrium point e.

Its Criticisms

The analysis of the cobweb theory is based upon very restrictive assumptions which make its applicability doubtful.

1. Not Realistic. It is not realistic to assume that the demand and supply conditions remain unchanged over the previous and current periods so that the demand and supply curves do not change (or shift). In reality, they are bound to change with considerable divergences between the actual and expected prices. Suppose the price is so low that some producers incur heavy losses. As a result, the number of sellers is reduced which changes the position of the supply curve. It is also possible that the expected price may be quite different from the estimated price. As a result, the cobweb may not develop properly on the basis of unchanged demand and supply curves. Thus demand, supply and price relations that lead to different cobwebs have little real applicability.

2. Output not Determined by Price. The thoery assumes that the output is determined by the price only. In reality, the agricultural output in particular is determined by several other factors also, such as weather, seeds, fertilizer, technology, etc.

3. Divergent Cobweb Impossible. Critics hold that divergent cobweb is impossible. It is obvious from Fig. 4 that once the equilibrium is upset, the cobweb cycle goes on diverging for an indefinite period which leads to an explosive situation. It is impossible.

4. Continuous Cobweb Impractical. Critics point out that continuous cobweb is impractical because it cannot continue indefinitely. This is because producers incur more loss than profit from it. This is explained by Fig. 5. If a farmer produces OQ output, he receives total revenue $OQbP$ whereas his total cost is $OQaP_1$ and his net profit is $PbaP_1$. When the output is OQ_1, total revenue is OQ_1dP_1 while the total cost is OQ_1cP.

Thus, he incurs $P_1 dcP$ total loss. Hence, in the case of continuous cobweb cycle, the producers have to face alternative years of profit and loss, but losses always exceed profit. Therefore, this cycle is impractical.

5. Not a Theory. In reality, cobweb is not a business cycle theory because it only explains fluctuations in the agricultural sector. So it is not used in explaining business cycles.

Implications

The cobweb model is an oversimplification of the real price determination process. But it supplies new information to the market participants about the market behaviour which they can incorporate into their decisions. The cobweb model is not merely an adjustment process of the market equilibrium but it also predicts unobservable events. Its significance lies in the demand, supply and price behaviour of agricultural commodities. Expectations about future conditions have an important influence on current prices. If there are boom conditions in the country, the farmers expect higher prices of their crops and increase their supplies in the market. But in the event of crop failures, the supplies of agricultural commodities will be reduced. In such a situation, the government may exempt farmers from agricultural taxes and even provide interest free loans to tide over the crisis. On the contrary, a bumper crop may lower the prices of agricultural crops by increasing their supplies more than their demand. In such a situation, the government may give subsidies to farmers or procure agricultural products at minimum support prices from the farmers.

6. Keynes's Theory

The Keynesian theory of the trade cycle is an integral part of his theory of income, output and employment. Trade cycles are periodic fluctuations of income, output and employment. Keynes regards the trade cycle as mainly due to "a cyclical change in the marginal efficiency of capital, though complicated and often aggravated by associated changes in the other significant short-period variables of the economic system."

According to Keynes, the principal cause of depression and

unemployment is the lack of aggregate demand. Revival can be brought about by raising aggregate demand which, in turn, can be raised by increasing consumption and/or investment. Since consumption is stable during the short-run, revival is possible by increasing investment. Similarly, the main cause of the downturn is reduction in investment. Thus in the Keynesian explanation of the trade cycle, "the cycle consists primarily of fluctuations in the rate of investment. And fluctuations in the rate of investment are caused mainly by fluctuations in the marginal efficiency of capital." The MEC (marginal efficiency of capital) depends on the supply price of capital assets and their prospective yield. Since the supply price of capital assets is stable in the short-run, the MEC is determined by the prospective yield of capital assets, which, in turn, depends on business expectations. Fluctuations in the rate of investment are also caused by fluctuations in the rate of interest. But Keynes gives more importance to fluctuations in the MEC as the principal cause of cyclical fluctuations.

To explain the course of the Keynesian cycle, we start with the *expansion phase*. During the expansion phase, the MEC is high. Businessmen are optimistic. There is rapid increase in the rate of investment. Consequently, output, employment and income increase. Every increase in investment leads to a multiple increase in income via the multiplier effect. This cumulative process of rising investment, income and employment continues till the boom is reached.

As the *boom* progresses, there is a tendency for the MEC to fall due to two reasons. *First,* as more capital goods are being produced steadily, the current yield on them declines. *Second,* at the same time the current costs of new capital goods rise due to shortages and bottlenecks of materials and labour.

During the *downturn,* investment falls due to a fall in the MEC and rise in the rate of interest. This leads to a cumulative decline in employment and income via the reverse operation of the multiplier. Further, the fall in the MEC may shift the consumption function downward thereby hastening the depression. Keynes attaches more importance to the sudden collapse of the MEC than to a rise in the rate of interest as an explanation of the

downturn of the cycle leading to the crisis and the depression.

Unlike the sudden collapse of the economic system, the *revival* takes time. It depends on factors which bring about the recovery of the MEC. "The time which must elapse before recovery begins, depends partly upon the magnitude of the *normal rate of growth of the economy* and partly upon *the length of life of capital goods*. The shorter the length of life of durable assets, the shorter the depression. And also, the more rapid the *rate of growth*, the shorter the depression." Another factor which governs the duration of depression is the "carrying costs of surplus stocks." According to Keynes, the carrying cost of surplus stocks during the depression is seldom less than 10 per cent per annum. So for a few years, disinvestment in stocks will continue till the surplus stocks are exhausted. Optimism takes the place of pessimism. The MEC increases. Fresh investment starts taking place. Revival has started.

Its Criticisms

Keynes's theory of the trade cycle is superior to the earlier theories because "it is more than a theory of the business cycle in the sense that it offers a general explanation of the level of employment, quite independently of the cyclical nature of changes in employment."[9] However, critics are not lacking in pointing out its weakness.

(1) Overemphasis on the Role of Expectations. Keynes has been criticised for his analysis of business cycle based on expectations. In fact, he overemphasised the role of expectations in influencing the MEC. According to Hart, Keynes relied on "convention" for forecasting changes in business expectations. The reliance on the conventional hypothesis makes Keynes' concept of expectations superfluous and unrealistic.

(2) Psychological Theory. Keynes considers the trade cycle as mainly due to fluctations in the MEC. The MEC, in turn, determines the rate of investment. And investment decisions, depend upon the psychology of businessmen or producers. Thus Keynes' theory is not much different from Pigou's psychological theory of the trade cycle.

(3) Explanation of Crisis Wrong. Keynes attributes the downturn to the

sudden collapse in the MEC. Accroding to Hazlitt, the term MEC being vague and ambiguous, "Keynes' explanation of the crisis of the marginal efficiency of capital is either a useless truism or an obvious error."

(4) Incomplete Theory. Another weakness of Keynes' theory of the trade cycle is that some of its variables such as expectations, MEC and investment cannot explain the different phases of the cycle. In the words of Dillard, "It is less than a complete theory of the business cycle because it makes no attempt to give a detailed account of the various phases of the cycle."

(5) Not Based on Empirical Data. Saulnier criticises Keynes's for lacking in factual proof. According to him, Keynes makes no attempt to test any of his deductions with facts. Dillard also points toward this defect when he writes that Keynes "does not examine closely the empirical data of cyclical fluctuations."

(6) One-Sided Theory. One of the serious omissions of Keynes's theory of the trade cycle is the acceleration principle. This made his theory one-sided because his explanation centres round the principle of multiplier. As pointed out by Sir John Hicks, "The theory of acceleration and the theory of multiplier are two sides of the theory of fluctuations, just as the theory of demand and the theory of supply are the two sides of the theory of value."

7. Samuelson's Model of Business Cycle

Prof. Samuelson[10] constructed a multiplier-accelerator model assuming one period lag and different values for the *MPC* (α) and the accelerator (β) that result in changes in the level of income pertaining to five different types of fluctuations.

The Samuelson model is

$$Y_t = G_t + C_t + I_t \qquad ...(1)$$

where Y_t is national income *Y* at time t which is the sum of government

expenditure G_t , consumption expenditure C_t and induced investment I_t.

$$C_t = \alpha Y_{t-1} \quad \text{...(2)}$$

$$I_t = \beta(C_t - C_{t-1}) \quad \text{...(3)}$$

Substituting equation (2) in (3) we have,

$$I_t = \beta(\alpha Y_{t-1} - \alpha Y_{t-2})$$

$$I_t = \beta\alpha Yt\text{-}1 - \beta\alpha Y_{t-2} \quad \text{...(4)}$$

$$G_t = 1 \quad \text{...(5)}$$

Substituting equations (2), (4) and (5) in (1) we have

$$Y_t = 1 + \alpha Y_{t-1} + \beta\alpha Y_{t-1} - \beta\alpha Y_{t-2} \quad \text{...(6)}$$

$$= 1 + \alpha(Y_{t-1} + \beta Y_{t-1}) - \beta\alpha Y_{t-2}$$

$$= 1 + \alpha(1+\beta) Y_{t-1} - \beta\alpha Y_{t-2} \quad \text{...(7)}$$

According to Samuelson, "If we know the national income for two periods, the national income for the following period can be simply derived by taking a weighted sum. The weights depend, of course, upon the values chosen for the marginal propensity to consume and for the *relation* (i.e.accelerator)". Assuming the value of the marginal propensity to consume to be greater than zero and less than one ($0 < \alpha < 1$) and of the accelerator greater than zero ($\beta > 0$), Samuelson explains five types of cyclical fluctuations which are summarised in the Table 1.

9. Dudley Dillard, *op.cit.*, p. 267.

Table 1. Samuelson's Interaction Model

Case	*Values*	*Behaviour of the Cycle*

1	$\alpha = .5, \beta = 0$	Cycleless Path
2	$\alpha = .5, \beta = 1$	Damped Fluctuations
3	$\alpha = .5, \beta = 2$	Fluctuations of Constant Amplitude
4	$\alpha = .5, \beta = 3$	Explosive Cycles
5	$\alpha = .5, \beta = 4$	Cycleless Explosive Path

Case 1 : Samuelson's case 1 shows a cycleless path because it is based only on the multiplier effect, the accelerator playing no part in it. This is shown in Fig. 6 (A).

Case 2 shows a damped cyclical path fluctuating around the static multiplier level and gradually subsiding to that level, as shown in Fig. 6 (B).

Case 3 depicts cycles of constant amplitude repeating themselves around the multiplier level. This case is depicted in Fig. 6 (C).

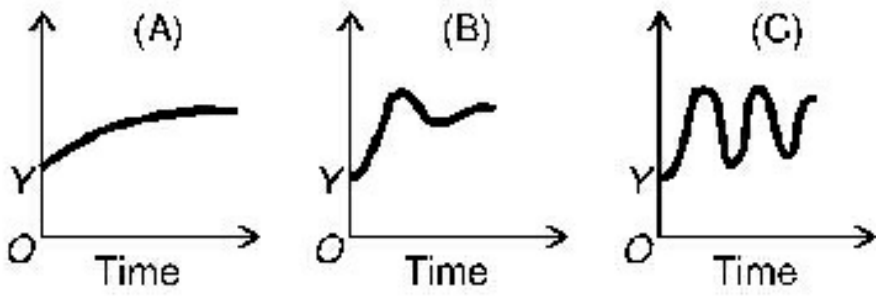
Case 4 reveals anti-damped or explosive cycles, see Fig. 6 (D).

Case 5 relates to a cycleless explosive upward path eventually approaching a compound interest rate of growth, as shown in Fig. 6 (E).

Of the five cases explained above, only three cases 2, 3 and 4 are cyclical in nature. But they can be reduced to two because case 3 pertaining to cycles of constant amplitude has not been experienced. So far as case 2 of damped cycles is concerned these cycles have been occurring irregularly in a milder form over last half century. Generally, cycles in the post-World War II period have been relatively damped compared to those in the inter-World War II period. They are the result of “such disturbances—which may be called erratic shocks—arising from exogenous factors, such as wars, changes in crops, inventions and so on ‘which’ might be expected to come alongwith fair persistence." But it is not possible to measure their magnitude.

Case 4 of explosive cycles has not been found in the past, its absence being the result of endogenous economic factors that limit the swings. Hicks has, however, built a model of the trade cycle assuming values that

would make for explosive cycles kept in check by ceilings and floors.



Critical Appraisal of the Model

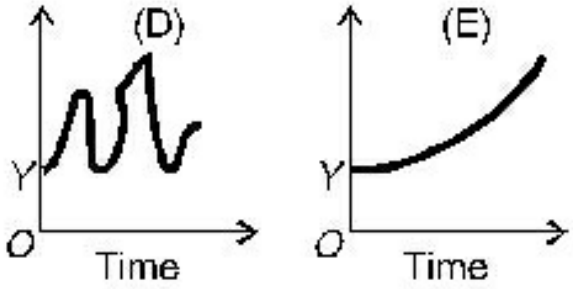


Fig. 6

The interaction of the multiplier and the accelerator has the merit of raising national income at a much faster rate than by either the multiplier or the accelerator alone. It serves as a useful tool not only for explaining business cycles but also as a guide to stabilisation policy. As pointed out by Prof. Kurihara, "It is in conjunction with the multiplier analysis based on the concept of marginal propensity to consume (being less than one) that the acceleration principle serves as a useful tool of business cycle analysis and a helpful guide to business cycle policy." The multiplier and the accelerator combined together produce cyclical fluctuations. The greater the value of the accelerator (β), the greater is the chance of an explosive cycle. The greater the value of the multiplier, the greater the chance of a cycleless path.

---

10. Paul A. Samuelson, "Interactions between the Multiplier Analysis and Principle of Acceleration," *R.E.S.*, May 1939.

** Consumption in period t is regarded as a function of the previous period (t–1).

Limitations

Despite these apparent uses of the multiplier-accelerator interaction, this analysis has its limitations:

(1) Samuelson is silent about the length of the period in the different cycles explained by him.

(2) This model assumes that the marginal propensity to consume (α) and the accelerator (β) are constants, but in reality they change with the level of income so that this is applicable only to the study of small fluctuations.

(3) The cycles explained in this model oscillate about a stationary level in a trendless economy. This is not realistic because an economy is not trendless but it is in a process of growth. This has led Hicks to formulate his theory of the trade cycle in a growing economy.

(4) According to Duesenberry, it presents a mechanical explanation of the trade cycle because it is based on the multiplier-accelerator interaction in rigid form.

(5) It ignores the effects of monetary changes upon business cycles.

8. Hicks's Theory of Business Cycle

J.R. Hicks in his book *A Contribution to the Theory of the Trade Cycle* builds his theory of business cycle around the principle of the multiplier-accelerator interaction. To him, "the theory of the acceleration and the theory of the multiplier are the two sides of the theory of fluctuations." Unlike Samuelson's model, it is concerned with the problem of *growth* and of a moving equilibrium.

Ingredients of the Theory

The ingredients of Hicks's theory of trade cycle are warranted rate of growth, consumption function, autonomous investment, an induced investment function, and multiplier-accelerator relation.

The warranted rate of growth is the rate which will sustain itself. It is consistent with saving-investment equilibrium. The economy is said to be growing at the warranted rate when real investment and real saving are taking place at the same rate. According to Hicks, it is the multiplier-accelerator interaction which weaves the path of economic fluctuations

around the warranted growth rate.

The consumption function takes the form $C_t = \alpha Y_{t-1}$. Consumption in period *t* is regarded as a function of income (*Y*) of the previous period (*t*–1). Thus consumption lags behind income, and the multiplier is treated as a lagged relation.

The autonomous investment is independent of changes in the level of output. Hence it is not related to the growth of the economy.

The induced investment, on the other hand, is dependent on changes in the level of output. Hence it is a function of the growth rate of the economy. In the Hicksian theory, the accelerator is based on induced investment which alongwith the multiplier brings about an *upturn*. The accelerator is defined by Hicks as the ratio of induced investment to the increase in income. Given *constant* values of the multiplier and the accelerator, it is the 'leverage effect' that is responsible for economic fluctuations.

Assumptions of the Theory

The Hicksian theory of trade cycle is based on the following assumptions:

(1) Hicks assumes a progressive economy in which autonomous investment increases at a constant rate so that the system remains in a moving equilibrium.

(2) The saving and investment coefficients are disturbed overtime in such a way that an upward displacement from equilibrium path leads to a lagged movement away from equilibrium.

(3) Hicks assumes *constant* values for the multiplier and the accelerator.

(4) The economy cannot expand beyond the full employment level of output. Thus "the full employment ceiling" acts as a direct restraint on the upward expansion of the economy.

(5) The working of the accelerator in the downswing provides an *indirect* restraint on the downward movement of the economy. The rate of

decrease in the accelerator is limited by the rate of depreciation in the downswing.

(6) The relation between the multiplier and accelerator is treated in a lagged manner, since consumption and induced investment are assumed to operate with a time lag.

(7) It is assumed that the average capital-output ratio (v) is greater than unity and that gross investment does not fall below zero. Thus the cycles are inherently explosive but are contained by ceilings and floors of the economy.

The Hicksian Theory

Hicks explains his theory of the trade cycle in terms of Fig. 7. Line *AA* shows the path of *autonomous* investment growing at a constant rate. *EE* is the equilibrium level of output which depends on *AA* and is deduced from it by the application of the multiplier accelerator interaction to it. Line *FF* is the full employment ceiling level above the equilibrium path *EE* and is growing at the constant rate of autonomous investment. *LL* is the lower equilibrium path of output representing the floor or 'slump equilibrium line'.

Hicks begins from a cycleless situation P_O on the equilibrium path *EE* when an increase in the rate of autonomous investment leads to an upward movement in income. As a result, the growth of output and income propelled by the combined operation of the multiplier and accelerator moves the economy on to the upward expansion path from P_0 to P_1. According to Hicks, this *upswing* phase relates to the *standard* cycle which will lead to an *explosive* situation because of the given values of the multiplier and the accelerator. But this does not happen because of the upper limit or ceiling set by the full employment level *FF*. Hicks writes in this connection: “I shall follow Keynes in assuming that there is some point at which output becomes inelastic in response to an increase in effective demand.” Thus certain bottlenecks of supply emerge which prevent output from reaching the peak and instead encounter the ceiling at P_1.

When the economy hits the full employment ceiling at P_1, it will creep along the ceiling for a period of time to P_2 and the downward swing will not start immediately. The economy will move along the ceiling from P_1 to P_2 depending upon the time period of the investment lag. The greater the investment lag, the more the economy will move along the ceiling path. Since income at this level is decreasing relative to the previous stage of the cycle, there is a decreased amount of investment. This much of investment is insufficient to keep the economy at the ceiling level, and then the downturn starts.

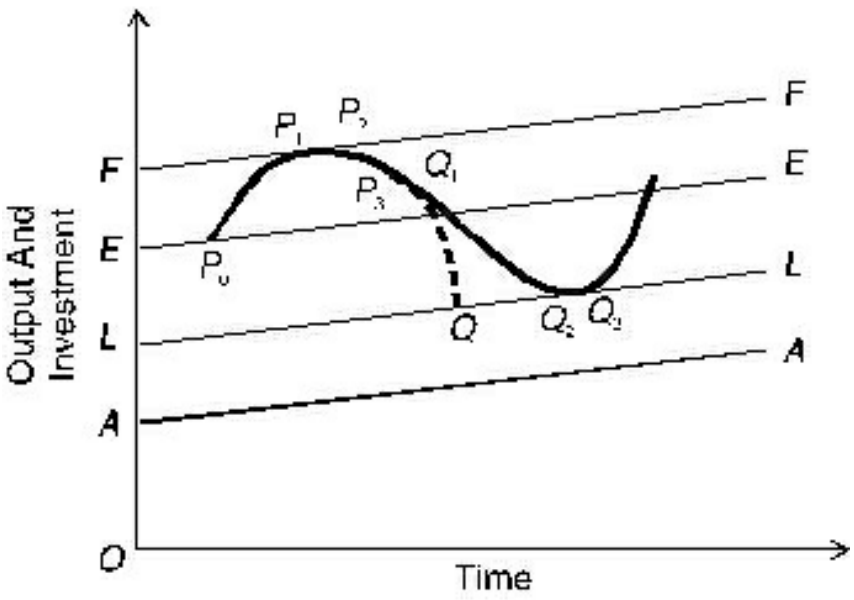


Fig. 7

During the *downswing*, "the multiplier-accelerator mechanism sets in reverse, falling investment reducing income, reduced income reducing investment, and so on, progressively. If the accelerator worked continuously, output would plunge downward below the equilibrium level *EE*, and because of explosive tendencies, to a greater extent than it rose above it." The fall in output in this case might be a steep one, as shown by P_2 P_3 Q. But in the downswing, the accelerator does not work so swiftly as in the upswing. If the slump is severe, induced investment will quickly fall to zero and the value of the accelerator becomes zero. The rate of decrease in investment is limited by the rate of depreciation. Thus the total amount of investment in the economy is equal to autonomous investment *minus* the constant rate of depreciation. Since autonomous investment is taking place, the fall in output is much gradual and the slump much longer than the boom, as indicated by Q_1Q_2. At Q_2, the slump reaches the bottom or floor provided by the *LL* line. The economy does not turn upward

immediately from Q_2 but will move along the slump equilibrium line to Q_3 because of the existence of excess capacity in the economy. Finally, when all excess capacity is exhausted, autonomous investment will cause income to rise which will in turn lead to an increase in induced investment so that the accelerator is triggered off which alongwith the multiplier moves the economy toward the ceiling again. It is in this way that the cyclical process will be repeated in the economy.

Its Criticisms

The Hicksian theory of the business cycle has been severely criticised by Duesenberry, Smithies and others on the following grounds:

1. Value of Multiplier not Constant. Hicks's model assumes that the value of the multiplier remains constant during the different phases of the trade cycle. This is based on the Keynesian stable consumption function. But this is not a realistic assumption, as Friedman has proved on the basis of empirical evidence that the marginal propensity to consume does not remain stable in relation to cyclical changes in income. Thus the value of the multiplier changes with different phases of the cycle.

2. Value of Accelerator not Constant. Hicks has also been criticised for assuming a constant value of the accelerator during the different phases of the cycle. The constancy of the accelerator presupposes a constant capital-output ratio. These are unrealistic assumptions because the capital-output ratio is itself subject to change due to technological factors, the nature and composition of investment, the gestation period of capital goods, etc. Lundberg, therefore, suggests that the assumption of constancy in accelerator should be abandoned for a realistic approach to the understanding of trade cycles.[11]

3. Autonomous Investment not Continuous. Hicks assumes that autonomous investment continues throughout the different phases of the cycle at a steady pace. This is unrealistic because financial crisis in a slump may reduce autonomous investment below its normal level. Further, it is also possible, as pointed out by Schumpeter, that autonomous investment may itself be subject to fluctuations due to a technological

innovation.

4. Growth not Dependent only on changes in Autonomous Investment. Another weakness of the Hicksian model is that growth is made dependent upon changes in autonomous investment. It is a burst of autonomous investment from the equilibrium path that leads to growth. According to Prof. Smithies,[12] the source of growth should lie within the system. In imputing growth to an unexplained extraneous factor, Hicks has failed to provide a complete explanation of the cycle.

5. Distinction Between Autonomous and Induced Investment not Feasible. Critics like Duesenberry[13] and Lundberg point out that Hicks's distinction between autonomous and induced investment is not feasible in practice. As pointed out by Lundberg, every investment is autonomous in the short run and a major amount of autonomous investment becomes induced in the long run. It is also possible that part of a particular investment may be autonomous and a part induced, as in the case of machinery. Hence this distinction between autonomous and induced investment is of doubtful validity in practice.

6. Ceiling Fails to Explain adequately the onset of Depression. Hicks has been criticised for his explanation of the ceiling or the upper limit of the cycle. According to Duesenberry, the ceilling fails to explain adequately the onset of depression. It may at best check growth and not cause a depression. Shortage of resources cannot bring a sudden decline in investment and thus cause a depression. The recession of 1953-54 in America was not caused by shortage of resources. Further, as admitted by Hicks himself, depression may start even before reaching the full employment ceiling due to monetary factors.

7. Explanation of Floor and Lower Turning Point not Convincing. Hicks’s explanation of the floor and of the lower turning point is not convincing. According to Hicks, it is autonomous investment that brings a gradual movement towards the floor and it is again increase in autonomous investment at the bottom that leads to the lower turning point. Harrod doubts the contention that autonomous investment would be increasing at the bottom of the depression. Depression may retard rather

than encourage autonomous investment. Further, Hicks's contention that revival would start with the exhaustion of excess capacity has not been proved by empirical evidence. Rendings Fels's study of the American business cycles in the 19th century has revealed that the revival was not due to the exhaustion of excess capacity. Rather in certain cases, revival started even when there was excess capacity.

8. Full Employment level not Independent of Output Path. Another criticism levelled against Hicks's model is that the full employment ceiling. As defined by Hicks, it is independent of the path of output. According to Dernburg and McDougall, the full employment level depends on the magnitude of the resources that are available to the country. The capital stock is one of the resources. When the capital stock is increasing during any period, the ceiling is raised. "Since the rate at which output increases determines the rate at which capital stock changes, the ceiling level of output will differ depending on the time path of output."

9. Explosive Cycle not Realistic. Hicks assumes in his model that the average capital-output ratio (v) is greater than unity for a time lag of one year or less. Thus explosive cycles are inherent in his model. But empirical evidence shows that the response of investment to a change in output (v) is spread over many periods. As a result, there have been damped cycles rather than explosive cycles.

10. Mechanical Explanation of Trade Cycle. Another serious limitation of the theory is that it presents a mechanical explanation of the trade cycle. This is because the theory is based on the multiplier-accelerator interaction in rigid form, according to Kaldor and Duesenberry. Thus it is a mechanical sort of explanation in which human judgement, business expectations and decisions play little or no part. Investment plays a leading role based on formula rather than on judgement.

11. Contraction Phase not Longer than Expansion Phase. Hicks has been criticised for asserting that the contraction phase is longer than expansion phase of trade cycle. But the actual behaviour of the postwar cycles has shown that the expansionary phase of the business cycle is much longer than the contractionary phase.

Conclusion. Despite these apparent weaknesses of the Hicksian model, it is superior to all the earlier theories in satisfactorily explaining the turning points of the business cycle. To conclude with Dernburg and McDougall, "The Hicks's model serves as a useful framework of analysis which, with modification, yields a fairly good picture of cyclical fluctuation within a framework of growth. It serves especially to emphasise that in a capitalist economy characterised by substantial amounts of durable equipment, a period of contraction inevitably follows expansion. Hicks's model also pinpoints the fact that in the absence of technical progress and other powerful growth factors, the economy will tend to languish in depression for long periods of time." The model is at best suggestive.

9. Goodwin's Trade Cycle Model

Goodwin[14] presents a non-linear model of trade cycle as against Hicks' linear model.

---

11. E. Lunberg, "*The Stability Growth*," *J.E.P.*, No. 8, 1958.

12. A Smithies, "Economic Fluctuations and Growth" *Econometrica*, Jan. 1957.

13. J.B. Duesenberry, *Business Cycles and Economic Growth*, 1958.

Features of the Model

The main features of Goodwin's model are :

(a) A linear consumption function $C = \alpha Y + C_0$ where α = consumption-income ratio, C_0=autonomous consumption.

(b) Net investment (I) is equal to the rate of change in capital stock which is the result of adjustment between actual capital, K and desired capital, $\overline{K}$ and $\overline{K} = \beta Y + \alpha$ where β is the accelerator.

(c) Desired capital $\overline{K}$ is proportional to output, Y.

(d) Net investment changes due to a change in output which in turn, changes the level of desired capital. As a result, the accelerator which is

non-linear brings a change in output via net investment.

Goodwin shows three possibilities for net investment :

(1) When $\overline{K} > K$, there will be increase in the rate of net investment which will shift the economy to full capacity output for capital goods industries.

(2) When $\overline{K} = K$, this situation is maintained by meeting the replacement demand and by keeping net investment as zero, *i.e.*, I = O.

(3) When $K > \overline{K}$, the adjustment takes place with the scrappage rate, K_2.

Cyclical Path

To explain the cyclical path of Goodwin's model, we start with $K = \overline{K}$ situation where the equilibrium level of output $Y = \frac{1}{1-\alpha}(C_0 + I)$ and $I = a$ is the technological growth factor. These take the system along an equilibrium path. This path is similar to the Hicksian line *EE*. The Hicksian path relates to a steady increase in autonomous investment while Goodwin's path relates to a rise in desired capital resulting from a continuous technical change. Any divergence from this equilibrium path will not bring the economy toward it and there will be continuous fluctuations around it.

Take $\overline{K} > K$, the desired capital stock being more than the actual capital, investment increases and the propelling forces of multiplier and accelerator will push the system in the upward direction. During this phase of expansion, given the constant rate of depreciation, net investment increases in proportion to the change in output. As the accelerator of Goodwin's model is non-linear, the increase in net investment is not related to the increase in output but to the difference between the actual

capital stock (K) and the desired capital stock ($\overline{K}$). This difference is determined by the rate of change in output. This investment relationship being continuous, investment is immediately pushed to the extreme position. Thus expansion reaches a situation where $I = \frac{dk}{dt} = K_1$, the rate of capital goods capacity output. This restricts the increase in output and the accelerator becomes discontinuous. During the expansion phase $\overline{K} > K$ but at the peak $K = \overline{K}$ which is due to a decline in the rate of autonomous investment.

Once the desired capital $\overline{K}$ exceeds the actual capital stock K, the desired capital $\overline{K}$ will be equal to the scrappage or replacement rate K_2. This leads to a fall in the desired capital $\overline{K}$ and in the rate of autonomous investment. Thus the contraction path of the system starts.

During the contraction phase when $K > \overline{K}$, gap between the two is to be met by the scrappage rate K_2. Simultaneously, the desired capital $\overline{K}$ continues to rise by the technological growth factor. The gap between K and $\overline{K}$ is being closed by the gradual elimination of excess capacity through failure to replace and the steady occurrence of capital-using innovations. In this way, the contraction path pushes the economic system to the lowest extreme point of $I = -K_2$.

When the economy reaches the lower turning point of $K = \overline{K}$, the level of desired capital $\overline{K}$ is to be raised. This is possible by stimulating the constant technical growth factor. When $\overline{K} > K$, the expansion again starts. So long as the technical progress continues, the economy cannot remain at the lower turning point. In fact, fast technical progress will expand the economy soon. As against, the Hicksian model, the Goodwin's cycle does not creep along the floor or the ceiling level. Rather, it jumps from these levels. Both boom and recession bring contraction and recovery of their own. If the growth factor does not exist in technological change, there will be no autonomous investment in the capital stock. Consequently, the

economy will remain stuck in the contraction phase for a long time instead of expanding. Because it takes more time for capital to depreciate than in capital formation. In such a situation there will be one-sided formation of trade cycles, as shown in Figure 8. In the beginning of the expansion phase, $\overline{K}$ increases to *OA* level and *K* also rises and equals it. When $K=\overline{K}$, contraction or downswing starts, $\overline{K}$ is reduced to *OB* level and *K* also declines and equals $\overline{K}$. There being no technological change in the growth factor, it takes more time for *K* to equal $\overline{K}$.

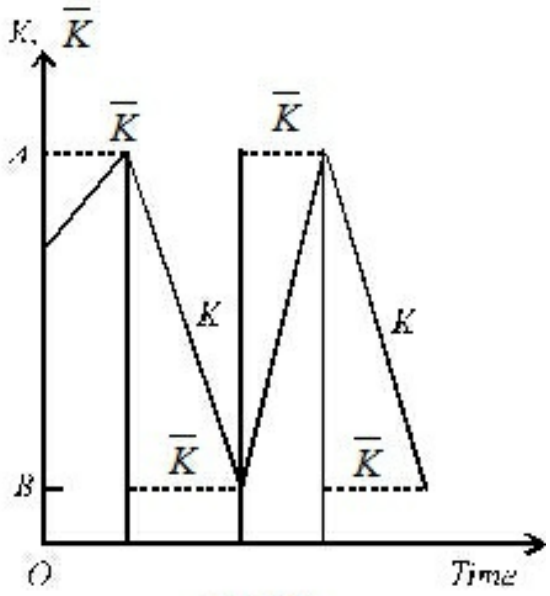


Fig. 8

There are two features of the Goodwin cycle in figure 8.

(1) In the expansion phase, $\overline{K}>K$ and in the contraction phase, $K>\overline{K}$.

(2) At the peak of each cycle, the common level of $\overline{K}$ is *OA* which is smaller than the *OB* level of $\overline{K}$ at the trough and shows more time for *K* to equal $\overline{K}$. This is clear from the dotted horizontal lines at *OA* level in the upper portion and the dotted horizontal lines at *OB* level in the lower portion of Figure 8.

On the other hand, if the growth factor for technological progress is included, $\overline{K}$ continues to increase with each cycle and it takes long time for *K* to equal $\overline{K}$ in the expansion phase. When the downswing starts, there being no growth factor the contraction phase becomes shorter. Thus the presence and absence of technological growth factor make the subsequent peaks and troughs at the higher level than the earlier peaks and troughs in the Goodwin cycle. These are shown as upper and lower dotted lines in Figure 9.

---

14. R.M. Goodwin, "The Non-Linear Accelerator and the Persistence of Business Cycles", *Econometrica*, Jan. 1951.

In the figure *aa* line shows the technological growth factor which is similar to the Hicksian *EE* line where $K=\bar{K}=a$. The upper and lower limits of $\bar{K}$ have been shown by the dotted lines. *P* represents peaks and *T* as troughs where $K=\bar{K}$. The contraction paths *PT* become shorter than the expansion paths, *TP*.

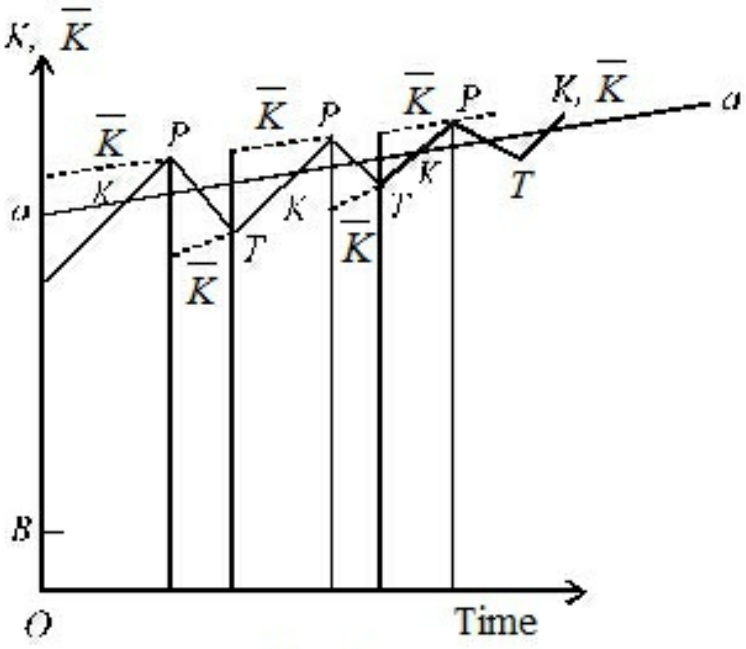


Fig. 9

Its Criticisms

The following are the weaknesses of Goodwin's model.

1.Unrealistic. The Goodwin model is unrealistic on two counts. *First*, it shows that the recessionary phase is longer than the expansionary phase of the cycle, as shown in Figure 9. *Second*, the desired stock of capital becomes constant at the peak, as shown by the dotted lines in Figure 9.

2. Empirically Wrong. The Goodwin model shows that when the economy reaches the peak and the trough, it at once takes a downturn and upturn respectively. This is empirically wrong because both recession and recovery are slow processes. As Hicks has also pointed out that the economy creeps along the upper and lower ceilings.

Conclusion. Despite these limitations, according to Prof. Allen, "The advantage of Goodwin's model is that the non-linear element is built in; the resulting oscillation maintains itself without any dependence on outside factors or on particular initial (or historical) conditions."[15]

Differences between Goodwin and Hicks Models

Goodwin's model is different from Hicks' model on the following points :

1. The Hicksian model shows only how cycles take place. But Goodwin's model traces the time path of a realistic cycle.

2. The Hicks model is linear which becomes non-linear when the ceiling and floor of a cycle meet. But Goodwin's model is throughout non-linear.

3. Hicks combines growth and cycles in his model and keeps growth dependant on autonomous investment. According to Goodwin, growth depends on technological change which continuously increases the productive capacity of the economy.

4. The equilibrium path *EE* of Hicks is based on continuous increase in autonomous investment. But Goodwin's growth path *aa* is based on the increase in desired capital $\overline{K}$ which is the result of continuous technological changes.

10. Friedman's Theory of Business Cycles*

Friedman and Schwartz[16] have argued on the basis of US historical data that business cycles are mostly monetary in origin. It is the money stock itself that shows a consistent cyclical behaviour which is closely related to the cyclical movements in economic activity at large. About the causal relation between the money stock and economic activity, they make the following generalisations: (i) changes in economic activity have always been accompanied by changes in the money stock; (ii) there have not been major changes in the money stock that have not been accompanied by changes in economic activity; and (iii) changes in the stock of money have been attributed to a specific variety of exogenous factors rather than to changes in economic activity. Thus changes in the money stock are a consequence as well as independent cause of changes in economic activity. But once they occur, they will, in turn, produce still further effects on economic activity. There is also much evidence that during business cycles the money stock plays largely an independent role.

In one of his earlier writings, Friedman emphasises that the concept of lag is related to the business cycle. These cycles are mostly monetary in origin. A monetary change effects different economic magnitudes, some of which adjust faster than others which cause distortions in economis

activity, thereby giving rise to the business cycles. Thus it can be said in Fisher's words that the cycle is largely a "dance of the dollar".

Major US historical economic fluctuations include inflationary and deep depression cycles. According to Friedman and Schwartz, the empirical evidence justifies the generalisations noted above. These show that the stock of money has displayed a systematic cyclical pattern over the decades. The money stock generally reaches its peak before the 'reference' peak of the cycles. Similarly, it reaches its trough before the 'reference' trough. One of their estimates of the lag between turning points in the growth of the money stock and in the level of economic activity reveals that during the seven cycles between 1927 and 1970, peaks in the rate of change in the money stock precede reference cycle peaks (in economic activity series) before downturns by 20 months on an average, and troughs in the rate of change of the money stock precede reference troughs by about 11 months on an average before upturns. The lag of economic activity appears to be greater for peaks than for troughs. There has been strong secular changes in the money stock over these decades. On the other hand, in deep depression cycles, there has been a greater fall in money stock. But in mild depressions, there has been a reduction in the growth rate of the money stock rather than any actual fall.

The usual cycle consists of a contraction phase in which economic activity declines to trough of the cycle, followed by expansion and reaching the peak of the cycle. These cycles are superimposed over a long run secular growth path, *GP*, as shown in Figure 10.

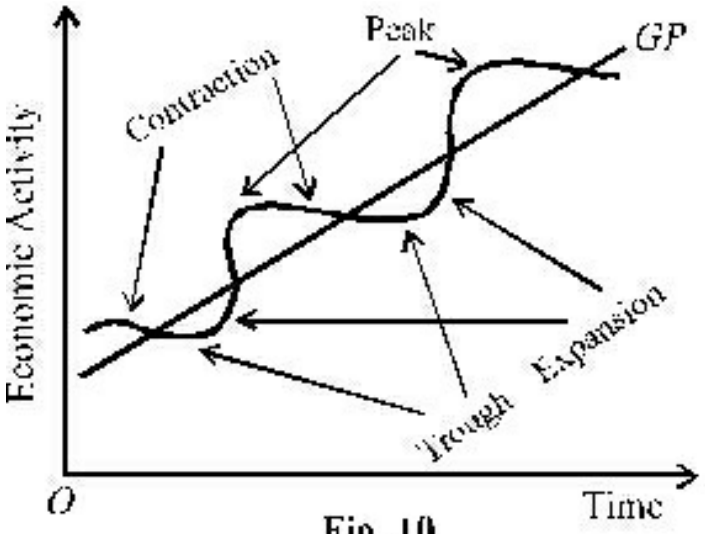


Fig. 10

On the basis of the above analysis, Friedman and Schwartz point toward two propositions: *First*, appreciable changes in the growth rate of the money stock are necessary and sufficient conditions for

appreciable changes in growth rate of economic activity or money income. *Second*, this is true both for long secular changes and also for changes over periods roughly the length of business cycles.

---

15. R.G.D. Allen, *Mathematical Economics*, 1969

They further observe that a *secular change* in the gowth rate of the money stock leading to longer period changes in money income are reflected mainly in different price behaviour rather than in different growth rates of output. On the other hand, a short period change in the growth rate of money stock also exerts a substantial influence on the growth rate of *output*. The greater stability of the "money multiplier" in contrast to the Keynesian investment multiplier has led Friedman and Schwartz to come to the above conclusion. According to them, substantial expansions in the quantity of money over short periods have been a major proximate source of the accompanying inflation of prices. Substantial contractions in the quantity of money over shorter periods have been a major factor in producing severe economic contractions. Cyclical variations in the quantity of money may well be an important element in the ordinary mild business cycle.

Next, Friedman and Schwartz explain the *mechanism* which brings about monetary changes leading to the business cycles. They begin their explanation of the transmission mechnaism with a state of moving equilibrium in which real per capita income, the stock of money, and the price level are changing at constant annual rates. Suppose the central bank increases the stock of money in the market by open market operations by purchasing securities. The non-bank sellers and commercial banks will try to readjust their portfolios. The commercial banks will create more money with increase in their reserves, thereby transmitting the increase in high-powered stock of money. On the other hand, the non-bank holders of cash will seek to purchase other categories of securities such as high-risk fixed coupons, equities, real property, etc. This will bid up prices of such assets. As the process continues, the initial impacts will spread throughout the economy. The increased demand for assets will spread sooner or later affecting equities, houses, durable producer goods, durable consumer goods, etc. All these will bid up the prices of assets and of both producer

and consumer goods. People will tend to consume more services, such as renting houses rather than purchasing them. This will tend to raise service prices. These effects will raise interest rates on the whole range of assets. Ultimately, expenditures rise on all directions without any change in interest rates at all. "Interest rates and asset prices may simply be conduit through which monetary change is transmitted to expenditures without being altered at all, just as a greater inflow into a tank may, after an interval, simply increase the rate of outflow without altering the level of the tank itself." All these forces operate simultaneously and there are cyclical fluctuations.

This explanation of the transmission mechanism fits with the empirical observations of business cycles. Initially, the rise in the growth rate of the money stock occurs early in the contraction phase. Its first impact is on the financial markets where first bonds, then equities and only later on payments for real resources will be affected. The financial markets tend to revive well before the trough. This is what has happened historically.

---

* This theory is also known as *Money and Business Cycles Theory*.

16. M. Friedman and A.W.Schwartz, "Money and Business Cycles", *R.E.S.*, (Supplement), 1963.

The dynamic process of transition from one equilibrium path to another involves a cyclical adjustment process. *Exogenous fluctuations* in the money stock will lead to fluctuations in the demand for goods and services. In addition, there may be an endogenous cycle. A rise in demand raises prices. If there is a lag in the adjustment of real money balances to the new price level, the initial portfolio adjustment will tend to overshoot. The initial rise in demand will thus be followed by a fall in demand. The result will be a damped cycle.

According to Friedman, the *lag* plays an important role in business cycles. The amplitude of economic fluctuations depends: *First*, on the amplitude, time pattern, number and independence of the disturbances affecting the economic system. *Second*, on the reaction mechanism of the economic system to the disturbances. The lag may be long because the effects of monetary disturbances are distributed over an extended period. A long lag may mean a larger damping of disturbances than a short lag. Hence there

is a smaller amplitude of resulting fluctuations. The lags in economic activity behind peaks and troughs in the rate of change of the money stock are not uniform. Friedman concludes on the basis of empirical evidence that lags involving changes in the rate of the money stock that affect the level of economic activity are both long and variable.

Its Criticisms

Economists have criticised Friedman's theory of money and business cycles on the following grounds:

(1) Monetary Changes not the Only Cause of Changes in Economic Activity. Friedman argues that it is monetary changes that cause changes in economic activity. But critics point out that the direction of causation is just the opposite of it. It is changes in the level of business activity which cause changes in the growth rate of the money stock.

(2) Monetary Changes not the Main Cause of Business Cycles. According to this theory, monetary changes are the main cause of business cycles. But in reality, business cycles are the result of the other exogenous factors like innovations. Monetary changes may be one among other factors, and not the only factor.

(3) Time Lag of Peaks and Troughs not Long and Variable. According to Friedman, the time lag of peaks and troughs in the rate of change of the money stock relative to economic changes in business cycles is both long and variable. Prof. Culbertson regards this evidence as faulty for two reasons: *First*, it relates turning points in one series in the money stock to turning points in economic activity. *Second*, it implies that monetary change has been an exogenous variable and that causation runs only from monetary change to economic change. In fact, causation also has run in other direction.

Despite these criticisms, it cannot be denied that one of the important causes of business cycles is "a dance of the dollar."

11. Kaldor's Model of the Trade Cycle

Nicholas Kaldor[17] built a model of the trade cycle based on the Keynesian terminology of saving and investment. He showed that the cycle is the result of pressures that push the economy toward the equality of ex-ante (anticipated, expected or planned) saving and investment. In fact, it is the difference between ex-ante saving and investment that leads to a cycle.

Kaldor shows the stability and

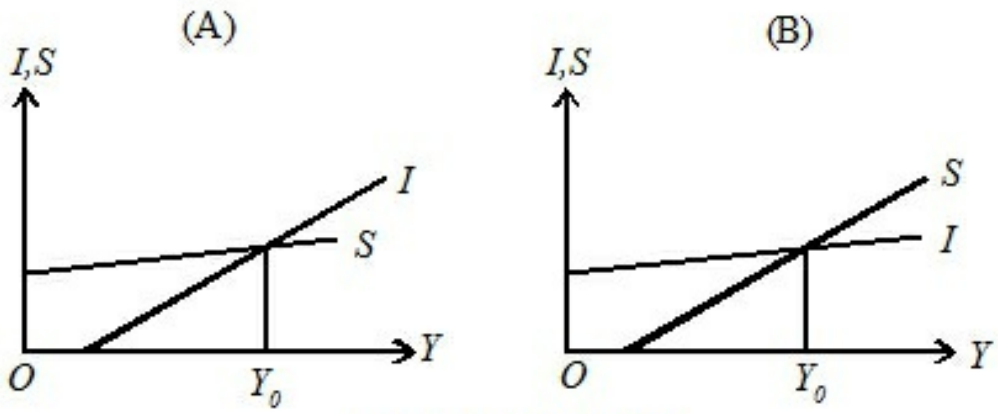


Fig. 11 (A) and (B)

instability conditions in the form of linear diagrams, though the cycle is only possible when *I* and *S* are non-linear. Take Figure 11 (A) and (B) where *I* and *S* are equal at the equilibrium level of income Y_0 . But in each case there is a single equilibrium position. In Panel (A) of the figure where *I*>*S* there is an unstable equilibrium position beyond Y_0 because such a situation will lead to limitless expansion, to full employment and hyper-inflation. On the other hand, if *S*>*I*, it means a downward movement to the left of Y_0 which will lead to zero output and employment and to collapse of the economy as shown in Panel (B) of the figure. Kaldor discards linear saving and investment functions because they fail to produce a cycle. Instead, be adopts non-linear saving and investment functions.

A non-linear investment function *I* is shown in Figure 12. As the economy moves into the expansion phase, shown by the movement from the left along the *I* curve, the curve is almost flat. It means that there is excess capacity at a low level of income and the net investment is zero. But "when expansion gets under way, the negative effect of accumulated capital is a more powerful influence for investment decisions than the higher levels of output and profit. In the opposite case of a high level of

income when the economy moves into the contraction phase, the *I* curve is again flat and the net investment is small "because rising costs of construction, increasing costs and increasing difficulty of borrowing will dissuade entrepreneurs from expanding still faster." This slows down the rate of increase in output. It means that the existing capital stock and the capacity are more than the current output. This leads to decline in further investment. Thus income falls and the cumulative effect is that the economy moves into the contraction phase.

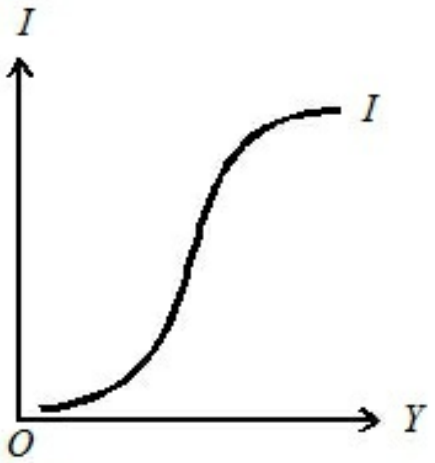


Fig. 12

Similarly, a non-linear saving function is shown in Figure 13. At very low levels of income, saving is much reduced and may even be negative. So during the expansion phase, the *MPS* is large. At normal levels of income, saving will increase at a smaller rate. This is shown by the middle range of the *S* curve. But at very high income levels, saving will be absolutely large and people will save a large proportion of their income.

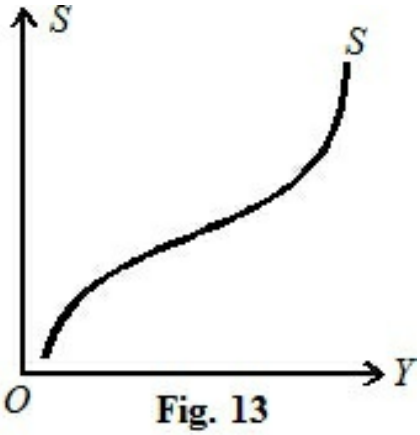


Fig. 13

The cycle is visible when the non-linear saving and investment curves are brought together, as in Figure 14. The figure shows multiple equilibria at positions *A, B* and *C.* Of these, *A* and *B* are stable positions and *C* an unstable position. Between positions *C* and *B* and below position *A, I* is greater than *S.* This will lead to the rise in income level. Between positions *A* and *C* and above *B* position, *S* is greater than *I.* This will lead to the fall in income.

But *A* and *B* are stable positions only in the short run. It is only in the long run that they become unstable and the path of the cycle is visible. For this, Kaldor introduces the capital stock as another variable that affects the relationship between saving and investment. He takes both saving and

investment as functions of income and capital stock so that

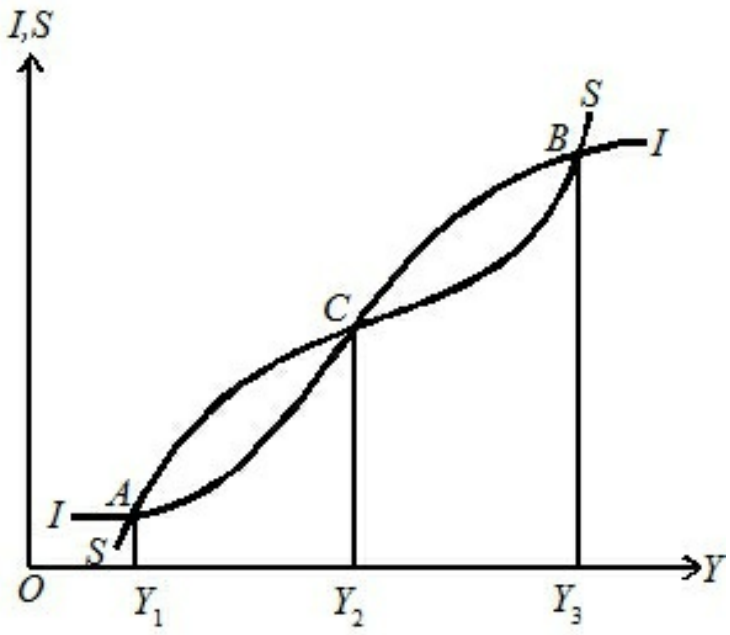


Fig. 14

$$S = f(Y, K)$$
$$I = f(Y, K)$$

$$\text{and} \quad \frac{dS}{dY} > 0 \quad , \quad \frac{dS}{dK} > 0$$

$$\frac{dI}{dY} > 0 \quad , \quad \frac{dI}{dK} < 0$$

and $\frac{dI}{dY} > \frac{dS}{dY}$, that is *MPI* is greater than

MPS over the expansion or contraction phase of the cycle.

The above relationships show that both *S* and *I* vary positively with *Y;* while *S v*aries directly with *K,* and *I* varies inversely with *K.* The relationship *MPI* > *MPS* shows the instability of the economy which will move it either toward expansion or contraction. In terms of Figure14, positions *A* and *B* are *"switch points"* in the long run. They are the points at which the economy alters its direction either toward expansion or contraction. Point *C* is unstable in both directions. It is only when points *C* and *B* come closer, the expansion phase of the cycle starts. When they are joined, expansion stops and contraction begins. On the contrary, when *C*

and *A* come closer contraction starts. When they are joined, contraction stops and expansion begins.

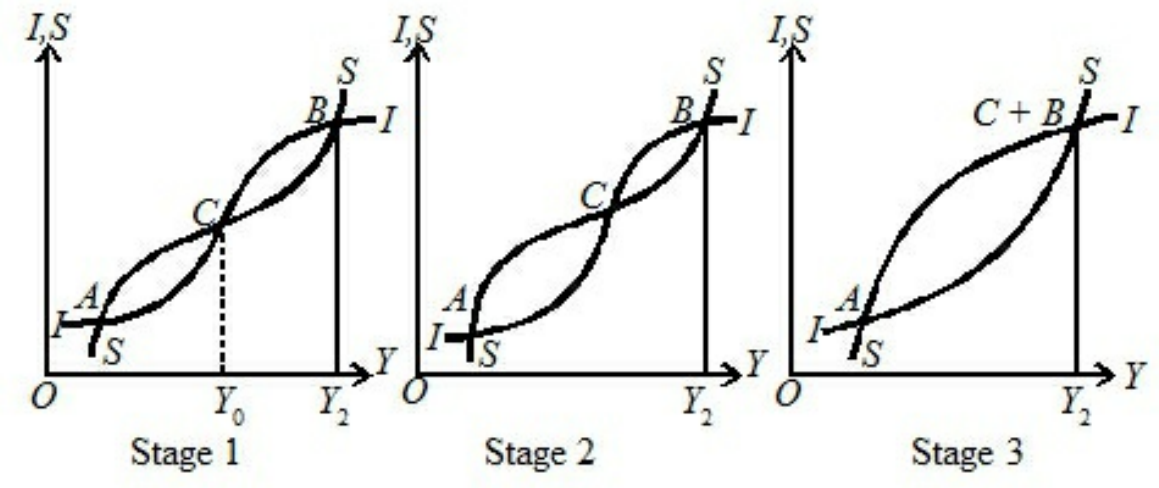


Fig. 15

Expansion Phase. Kaldor shows the expansion phase of his trade cycle in three stages, as shown in Figure 15. Starting from position Y_0 in figure stage 1 (which is the same as Figure 13), suppose the economy is in equilibrium at point C. But this is the point of unstable equilibrium. An upward displacement shows that *I*>*S* which leads to the economy towards the expansion path. As the rate of invesment is high, the economy's capital stock increases at a rapid rate. But as the capital stock increases, the MEC declines and investment curve shifts downward At the same time as the economy's capital stock increases, it raises the income of the economy thereby raising its saving. Thus the saving curve shifts upward. So a downward shift of the investment curve *I* and upward shift of the saving curve *S* bring the point *C* nearer to *B,* as shown in figure stage 2. This process of the downward shifting of the *I* curve and upward shifting of the *S* curve continues till the two curves are tangential and points *C* and *B* coincide, as shown in figure stage 3. But at this position *S*>*I* in both directions. So this is an unstable equilibrium position in the downward direction. This leads to the downward movement of the economy till point *A* is reached in stage 3.

---

17. "A Model of the Trade Cycle", *E.J.*, March, 1940.

Contraction Phase. The contraction of the trade cycle is also shown in three stages, as in Figure 16. We start from position Y_1 corresponding to point *A* in stage 4 of the figure. It is the point of short-run stable equilibrium but at a very low income level. But over the long run at such a low level of income, the capital stock decreases due to excess capacity and the investment curve *I* shifts upward. Simultaneously, saving falls which shifts the saving curve downward. Thus the shifting of the *I* curve upward and of the *S* curve downward bring positions *A* and *C* nearer, as shown in stage 5. This process will continue gradually till *I* and *S* curves arc tangential and positions *A* and *C* coincide, as in stage *6* figure. But this *A+C* position at Y_1 income level is unstable in the upward direction because *I>S.* This will lead to an expansionary process till the economy reaches the higher level of income Y_2 at point *B.* From *B,* the *I* and *S* curves gradually reach the positions shown in stage 1 of Figure 15, and again the cyclical process starts. Thus Kaldor's cyclical process is self-generating.

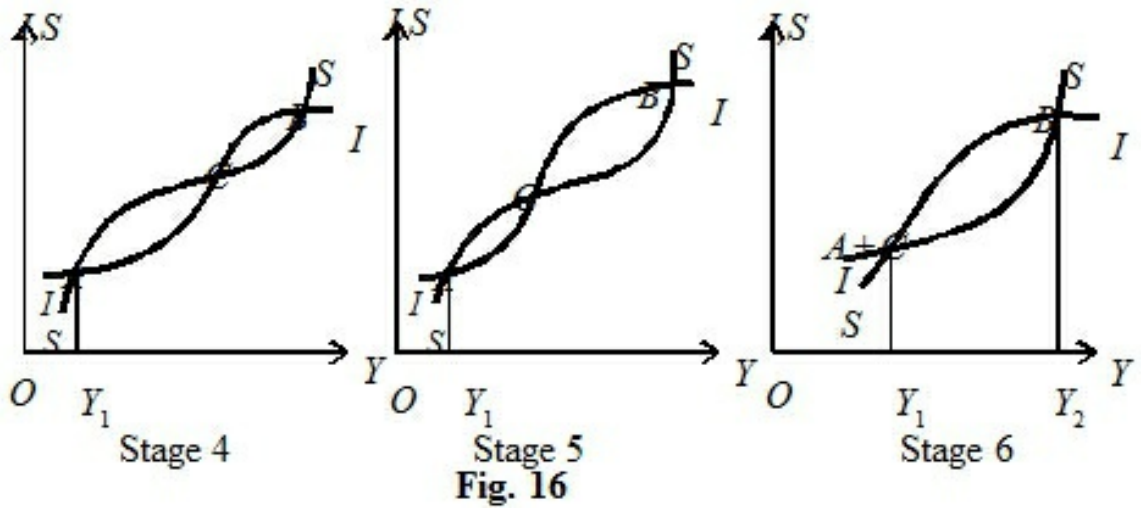


Fig. 16

According to Kaldor, the forces which bring about the lower turning point are not so certain at the higher level. "A boom left to itself, is certain to come to an end; but the depression might get into a position of stationariness, and remain there until external changes (the discovery of new inventions or of the opening up of new markets) come to the rescue."

Further, cycles in the Kaldor model are not necessarily of the same length and duration. Neither are expansions and contractions necessarily

symmetrical. In fact, these depend upon the slopes of the *I* and *S* curves and the rate at which they shift in each phase of the cycle.

Kaldor neither uses the acceleration principle nor the monetary factors in explaining his theory of the trade cycle. At the same time, he demonstrates how a cycle could exist in the absence of any growth factor.

MEASURES TO CONTROL BUSINESS CYCLES OR STABILISATION POLICIES

Various measures have been suggested and put into practice from time to time to control fluctuations in an economy. They aim at stabilising economic activity so as to avoid the ill-effects of a boom and a depression. The following three measures are adopted for this purpose.

1. Monetary Policy

Monetary policy as a method to control business fluctuations is operated by the central bank of a country. The central bank adopts a number of methods to control the quantity and quality of credit. To control the expansion of money supply during a *boom*, it raises its bank rate, sells securities in the open market, raises the reserve ratio, and adopts a number of selective credit control measures such as raising margin requirements and regulating consumer credit. Thus the central bank adopts a *dear money policy*. Borrowings by business and trade become dearer, difficult and selective. Efforts are made to control excess money supply in the economy.

To control a *recession or depression*, the central bank follows an *easy* or *cheap* monetary policy by increasing the reserves of commercial banks. It reduces the bank rate and interest rates of banks. It buys securities in the open market. It lowers margin requirements on loans and encourages banks to lend more to consumers, businessmen, traders, etc.

Limitations of Monetary Policy

But monetary policy is not so effective as to control a boom and a

depression. If the boom is due to cost-push factors, it may not be effective in controlling inflation, aggregate demand, output, income and employment. So far as depression is concerned, the experience of the Great Depression of 1930s tells us that when there is pessimism among businessmen, the success of monetary policy is practically nil. In such a situation, they do not have any inclination to borrow even when the interest rate is very low. Similarly, consumers who are faced with reduced incomes and unemployment cut down their consumption expenditure. Neither the central bank nor the commercial banks are able to induce businessmen and consumers to raise the aggregate demand. Thus the success of monetary policy to control economic fluctuations is severely limited.

2. Fiscal Policy

Monetary policy alone is not capable of controlling business cycles. It should, therefore, be supplemented by compensatory fiscal policy. Fiscal measures are highly effective for controlling excessive government expenditure, personal consumption expenditure, and private and public investment during a boom. On the other hand, they help in increasing government expenditure, personal consumption expenditure and private and public investment during a depression.

Policy during Boom. The following measures are adopted during a boom. During a boom, the government tries to *reduce unnecessary expenditure* on non-development activities in order to reduce its demand for goods and services. This also puts a check on private expenditure which is dependent on the government demand for goods and services. But it is difficult to cut government expenditure. Moreover, it is not possible to distinguish between essential and non-essential government expenditure. Therefore, this measure is supplemented by taxation.

To cut personal expenditure, the government *raises* the *rates* of personal, corporate and commodity taxes.

The government also follows the policy of having a *surplus budget* when the government revenues exceed expenditures. This is done by increasing the tax rates or reduction in government expenditure or both. This tends to

reduce income and aggregate demand through the reverse operation of the multiplier.

Another fiscal measure which is usually adopted is to *borrow* more from the public which has the effect of reducing the money supply with the public. Further, the repayment of public debt should be stopped and postponed to some future date when the economy stabilises.

Policy during Depression. During a depression, the government increases public expenditure, reduces taxes and adopts a budget deficit policy. These measures tend to raise aggregate demand, output, income, employment and prices. An increase in public expenditure increases the aggregate demand for goods and services and leads to increase in income via the multiplier. The public expenditure is made on such public works as roads, canals, dams, parks, schools, hospitals and other construction works. They create demand for labour and the products of private construction industries and helps in reviving them. The government also increases its expenditure on such relief measures as unemployment insurance, and other social security measures in order to stimulate the demand for consumer goods industries. Borrowing by the government to finance budget deficits utilises idle money lying with the banks and financial institutions for investment purposes.

Conclusion. The effectiveness of anti-cyclical fiscal policy depends upon proper timing of policy action and the nature and volume of public works and their planning.

3. Direct Controls

The aim of direct controls is to ensure proper allocation of resources for the purpose of price stability. They are meant to affect strategic points of the economy. They affect particular consumers and producers. They are in the form of rationing, licensing, price and wage controls, export duties, exchange controls, quotas, monopoly control, etc. They are more effective in overcoming bottlenecks and shortages arising from inflationary pressures. Their success depends on the existence of an efficient and honest administration. Otherwise, they lead to black marketing, corruption, long queues, speculation, etc. Therefore, they should be

resorted to only in emergencies like war, crop failures and hyper-inflation.

Conclusions

Of the various instruments of stabilisation policy, no single method is sufficient to control cyclical fluctuations. Therefore, all methods should be used simultaneously. This is because monetary policy is easy to apply but less effective while fiscal measures and direct controls are difficult to operate but are more effective. Since cyclical fluctuations are inherent in the capitalist system, they cannot be eliminated completely. Some fluctuations may be beneficial for economic growth and others may be undesirable. Stabilisation policy should, therefore, control undesirable fluctuations. We conclude with Keynes, “The right remedy for the trade cycles is not to be found in abolishing booms and thus keeping us permanently in a semi-slump; but in abolishing slumps and thus keeping us permanently in a quasi-boom.”

SELECTED READINGS

J.A.Estey, *Business Cycles*.

G.Haberler, *Prosperity and Depression*.

J.A.Schumpeter, *Business Cycles*, 1939.

D. Hamberg, *Business Cycles*, 1937.

S. Bober, The *Economics of Cycles and Growth*, 1968.

EXERCISES

1. What do you mean by a business cycle ? Explain its various phases.
2. Give the meaning, phases and effects of a business cycle.
3. Discuss critically the view that "the trade cycle is purely monetary phenomenon."
4. Critically examine Schumpeter's theory of innovations as the explanation of cyclical fluctuations.

5. Discuss the view that trade cycle is mainly due to fluctuations in the marginal efficiency of capital.
6. Discuss critically Samuelson's model of trade cycles.
7. Show that the interaction of the multiplier and accelerator provides a satisfactory explanation of the trade cycle.
8. Critically discuss the Hicksian theory of the trade cycle.
9. Explain Kaldor's model of the trade cycle.
10. Discuss Goodwin's model of the trade cycle and compare it with Hicks model.
11. Discuss Friedman's monetary theory of the trade cycle.
12. What is a business cycle ? Explain its characteristics. How can a business cycle be controlled ?

PART-VI

GROWTH MODELS

CHAPTER 39

THE HARROD-DOMAR MODELS

INTRODUCTION

The Harrod-Domar models of economic growth are based on the experiences of advanced economies. They are primarily addressed to an advanced capitalist economy and attempt to analyse the requirements of steady growth in such economy.

REQUIREMENTS OF STEADY GROWTH

Both Harrod and Domar are interested in discovering the rate of income growth necessary for a smooth and uninterrupted working of the economy. Though their models differ in details, yet they arrive at similar conclusions.

Harrod and Domar assign a key role to investment in the process of economic growth. But they lay emphasis on the dual character of investment. *Firstly,* it creates income, and *secondly,* it augments the productive capacity of the economy by increasing its capital stock. The former may be regarded as the '*demand effect*' and the latter the '*supply effect*' of investment. Hence so long as net investment is taking place, real income and output will continue to expand. However, for maintaining a full employment equilibrium level of income from year to year, it is necessary that both real income and output should expand at the same rate at which the productive capacity of the capital stock is expanding. Otherwise, any divergence between the two will lead to excess of idle capacity, thus forcing entrepreneurs to curtail their investment

expenditures. Ultimately, it will adversely affect the economy by lowering incomes and employment in the subsequent periods and moving the economy off the equilibrium path of steady growth. Thus, if full employment is to be maintained in the long run, net investment should expand continuously. This further requires continuous growth in real income at a rate sufficient enough to ensure full capacity use of a growing stock of capital. This required rate of income growth may be called the *warranted rate of growth* or "the full capacity growth rate."

Assumptions

The models constructed by Harrod and Domar are based on the following assumptions:

(1) There is an initial full employment equilibrium level of income.

(2) There is the absence of government interference.

(3) These models operate in a dosed economy which has no foreign trade.

(4) There are no lags in adjustments between investment and creation of productive capacity.

(5) The average propensity to save is equal to the marginal propensity to save.

(6) The marginal propensity to save remains constant.

(7) The capital coefficient, *i.e.*, the ratio of capital stock to income is assumed to be fixed.

(8) There is no depreciation of capital goods which are assumed to possess infinite life.

(9) Saving and investment relate to the income of the same year.

(10) The general price level is constant, *i.e.*, the money income and real income are the same.

(11) There are no changes in interest rates.

(12) There is a fixed proportion of capital and labour in the productive process.

(13) Fixed and circulating capitals are lumped together under capital

(14) There is only one type of product.

All these assumptions are not necessary for the final solution of the problem, nevertheless they serve the purpose of simplifying the analysis.

THE DOMAR MODEL

Domar[1] builds his model around the following question: since investment generates income on the one hand and increases productive capacity on the other at *what rate investment should increase in order to make the increase in income equal to the increase in productive capacity, so that full employment is maintained?*

He answers this question by forging a link between aggregate supply and aggregate demand through investment.

1. Evsey Domar, "Expansion and employment" in *A.E.R.*, March 1947, and *Essays on the Theory of Economic Growth*, 1957.

Increase in Productive Capacity. Domar explains the *supply* side like this. Let the annual rate of investment be I, and the annual productive capacity per dollar of newly created capital be equal on the average to s (which represents the ratio of increase in real income or output to an increase in capital or the reciprocal of the accelerator or the marginal capital-output ratio). Thus the productive capacity of I dollar invested will be $I.s$ dollars per year.

But some new investment will be at the expense of the old. It will, therefore, compete with the latter for labour markets and other factors of production. As a result, the output of old plants will be curtailed and the

increase in the annual output (productive capacity) of the economy will be somewhat less than *I.s.* This can be indicated as *I* σ, where σ (sigma) represents the net potential social *average productivity of investment* (= Δ *Y/I).* Accordingly *I*σ is less than *I.s. I*σ is the total net *potential increase in output* of the economy and is known as the *sigma* effect. In Domar's words, this "is the increase in output which the economy can produce," it is the "supply side of our system."

Required Increase in Aggregate Demand. The *demand* side is explained by the Keynesian multiplier. Let the annual increase in income be denoted by Δ*Y* and the increase in investment by Δ*I* and the propensity to save by α (alpha) (=Δ*S*/Δ *Y*). Then the increase in income will be equal to the multiplier (1/α) times the increase in investment:

$$\Delta Y = \Delta I \frac{1}{\alpha}$$

Equilibrium. To maintain full employment equilibrium level of income, aggregate demand should be equal to aggregate supply. Thus we arrive at the *fundamental equation* of the model:

$$\Delta I \frac{1}{\alpha} = I\alpha$$

Solving this equation by dividing both sides by *I* and multiplying by α we get:

$$\frac{\Delta I}{I} = \alpha\sigma$$

This equation shows that to maintain full employment the growth rate of *net autonomous investment* (Δ*I*/*I*) must be equal to ασ (the MPS times the productivity of capital). This is the rate at which investment must grow to assure the use of potential capacity in order to maintain a steady growth rate of the economy at full employment.

Domar gives a numerical example to explain his point: Let σ =25 per cent per year, α =12 per cent and *Y* =150 billion dollars per year. If full employment is to be maintained, an amount equal to 150 ×12/100 =18-

billion dollars should be invested. This will raise productive capacity by the amount invested σ times, *i.e.*, by 150 × 12/100 × 25/100 = 4.5 billion dollars, and the national income will have to rise by the same amount. But the relative rise in income will equal the absolute increase divided by the income itself, *i.e.*,

$$150 \times \frac{\frac{12}{100} \times \frac{25}{100}}{150} = \frac{12}{100} \times \frac{25}{100} = \alpha\sigma = 3 \text{ per cent}$$

Thus in order to maintain full employment, income must grow at the rate of 3 per cent per annum. This is the equilibrium rate of growth. Any divergence from this 'golden path' will lead to cyclical fluctuations. When $\Delta I/I$ is greater than ασ the economy would experience boom and when $\Delta I/I$ is less than ασ, it would suffer from depression.

THE HARROD MODEL

R.F. Harrod[2] tries to show in his model how steady *(i.e.*, equilibrium) growth may occur in the economy. Once the steady growth rate is interrupted and the economy falls into disequilibrium, cumulative forces tend to perpetuate this divergence thereby leading to either secular deflation or secular inflation.

The Harrod Model is based upon three distinct rates of growth. *Firstly*, there is the *actual growth rate* represented by G which is determined by the saving ratio and the capital-output ratio. It shows short-run cyclical variations in the rate of growth. *Secondly*, there is the *warranted growth rate* represented by *Gw* which is the full capacity growth rate of income of an economy. *Lastly*, there is the *natural growth rate* represented by *Gn* which is regarded as 'the welfare optimum' by Harrod. It may also be called the potential or the full employment rate of growth.

The Actual Growth Rate. In the Harrodian model the first fundamental equation is:

$$GC = s \qquad \ldots.(1)$$

where *G* is the rate of growth of output in a given period of time and can be expressed as Δ *Y/Y; C* is the net addition to capital and is defined as the ratio of investment to the increase in income, *i.e., I* /Δ *Y* and s is the average propensity to save, *i.e., SlY.* Substituting these ratios in the above equation we get:

$$\frac{\Delta Y}{Y} \times \frac{I}{\Delta Y} = \frac{S}{Y} \text{ or } \frac{I}{Y} = \frac{S}{Y} \text{ or } I = S$$

The equation is simply a re-statement of the truism that ex-post (actual, realized) savings equal ex-post investment.

The above relationship is disclosed by the behaviour of income. Whereas *S* depends on *Y, I* depends on the increment in income (Δ*Y*), the latter is nothing but the acceleration principle.

The Warranted Rate of Growth. The warranted rate of growth is, according to Harrod, the rate "at which producers will be content with what they are doing." It is the "entrepreneurial equilibrium; it is the line of advance which, if achieved, will satisfy profit takers that they have done the right thing." Thus this growth rate is primarily related to the behaviour of businessmen. At the warranted rate of growth, demand is high enough for businessmen to sell what they have produced and they will continue to produce at the same percentage rate of growth. Thus, it is the path on which the supply and demand for goods and services will remain in equilibrium, given the propensity to save. The equation for the warranted rate is

$$GwCr = s \qquad \text{...(2)}$$

where *Gw* is the warranted rate of growth or the full capacity rate of growth of income which will fully utilize a growing stock of capital that will satisfy the entrepreneurs with the amount of investment actually made. It is the *value* of Δ*Y/Y. Cr,* the capital requirements, denotes the amount of capital needed to maintain the warranted rate of growth, , *i.e.,* required capital-output ratio. It is the *value* of *I/* Δ *Y,* or *C. s* is the same as in the first equation, *i.e., S/Y.*

2. R.F. Harrod, *Towards a Dynamic Economics,* 1948.

The equation, therefore, states that if the economy is to advance at the steady rate of *Gw* that will fully utilize its capacity, income must grow at the rate of *s/Cr* per year; i.e., *Gw=s/Cr.*

If income grows at the warranted rate, the capital stock of the economy will be fully utilised and entrepreneurs will be willing to continue to invest the amount of saving generated at full potential income. *Gw* is therefore a self-sustaining rate of growth and if the economy continues to grow at this rate, it will follow the equilibrium path.

Genesis of Long-run Disequilibria. Full employment growth, the actual growth rate of *G* must equal *Gw,* the warranted rate of growth that would give steady advance to the economy, and *C* (the actual capital goods) must equal *Cr* (the required capital goods for steady growth).

If *G* and *Gw* are not equal, the economy will be in disequilibrium. For instance, if *G* exceeds *Gw,* then *C* will be less than *Cr.* When *G>Gw,* shortages result. There will be insufficient goods in the pipeline and/or insufficient equipment. Such a situation leads to *secular inflation* because actual income grows at a faster rate than that allowed by the growth in the productive capacity of the economy. It will further lead to a deficiency of capital goods, the acral amount of capital goods being less than the required capital goods *(C<Cr).* Under the circumstances, desired *(ex-ante)* investment would be greater than saving and aggregate production would fall short of aggregate demand. There would thus be *chronic inflation.* This is illustrated in Fig. 1 (A) where the growth rates of income are taken on the vertical axis and time on the horizontal axis. Starting from the initial full employment level of income Y_0, the actual growth rate *G* follows the warranted growth path *Gw* up to point *E* through period t_2. But from t_2 onward *G* deviates from *Gw* and is higher than the latter. In subsequent periods, the deviation between the two becomes larger and larger.

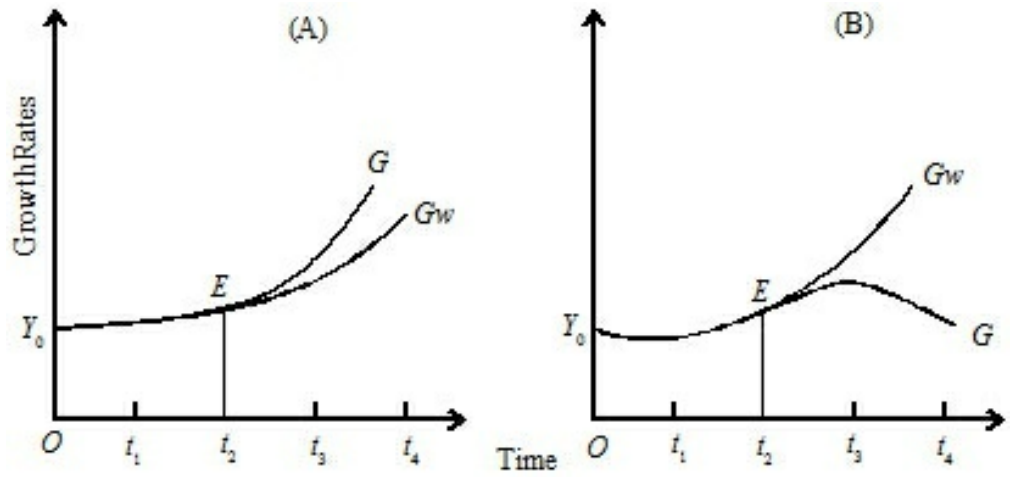


Fig. 1.

If, on the other hand, *G* is less that *Gw*, then *C* is greater than *Cr*. Such a situation leads to *secular depression* because actual income grows more slowly than what is required by the productive capacity of the economy leading to an excess of capital goods *(C>Cr).* This means that desired investment is less than saving and that the aggregate demand falls short of aggregate supply. The result is fall in output, employment, and income. There would thus be chronic depression. This is illustrated in Fig. 1 (B) when from period t_2 onward *G* falls below *Gw* and the two continue to deviate further away.

Harrod states that once *G* departs from *Gw*, it will depart further and further away from equilibrium. He writes: "Around that line of advance which if adhered to would alone give satisfaction centrifugal forces are at work, causing the system to depart further and further from the required line of advance." Thus the equilibrium between *G* and *Gw* is a *knife-edge equilibrium.* For once it is disturbed, it is not self-correcting. It follows that one of the major tasks of public policy is to bring *G* and *Gw* together in order to maintain long-run stability. For this purpose, Harrod introduces his third concept of the natural rate of growth.

The Natural Rate of Growth. The natural rate of growth is the rate of advance which the increase of population and technological improvements allow. It depends on the macro variables like population, technology, natural resources and capital equipment. In other words, it is the rate of increase in output at full employment as determined by a growing

population and the rate of technological progress. The equation for the natural rate of growth is

$$Gn.\ Cr= \text{ or } \#\ s$$

Here *Gn* is the *natural* or *full employment* rate of growth.

Divergence of *G, Gw* **and** *Gn.* Now for full employment equilibrium growth *Gn=Gw = G.* But this is a *knife-edge* balance. For once there is any divergence between natural, warranted and actual rates of growth conditions of secular stagnation or inflation would be generated in the economy. If *G>Gw,* investment increases faster than saving and income rises faster than *Gw.* If *G<Gw,* saving increases faster than investment and rise of income is less than *Gw.* Thus Harrod points out that if *Gw>Gn, secular stagnation* will develop. In such a situation, *Gw* is also greater than *G* because the upper limit to the actual rate is set by the natural rate as shown in Fig.2(A). When *Gw* exceeds *Gn, C > Cr* and there is an excess of capital goods due to a shortage of labour. The shortage of labour keeps the rate of increase in output to a level less than *Gw*. Machines become idle and there is excess capacity. This further dampens investment, output, employment and income. Thus the economy will be in the grip of chronic depression. Under such conditions saving is a *vice*.

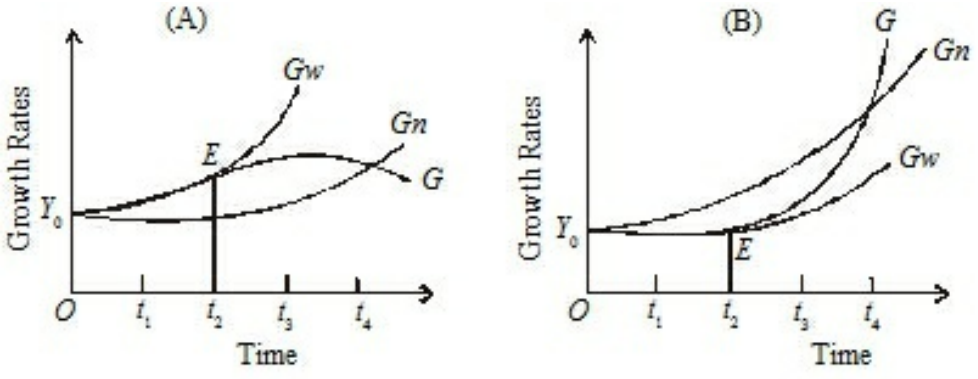


Fig. 2.

If *Gw<Gn, Gw* is also less than *G* as shown in Fig. 2(B). The tendency is for *secular inflation* to develop in the economy. When *Gw* is less than *Gn, C< Cr.* There is a shortage of capital goods and labour is plentiful. Profits are high since desired investment is greater than realised investment and

the businessmen have a tendency to increase their capital stock. This will lead to secular inflation. In such a situation saving is a *virtue* for it permits the warranted rate to increase.

This instability in Harrod's model is due to the rigidity of its basic assumptions. They are a fixed production function, a fixed saving ratio, and a fixed growth rate of labour force. Economists have attempted to relieve this rigidity by permitting capital and labour substitution in the production function, by making the saving ratio a function of the profit rate and the growth rate of labour force as a variable in the growth process.

The *policy implications* of the model are that saving is a virtue in any inflationary gap economy and vice in a deflationary gap economy. Thus in an advanced economy, *s* has to be moved up or down as the situation demands.

A Comparative Study of the Two Models

Points of Similarity. The following are the points of similarity in the two models.

The Domar Model		The Harrod Model	
$\sigma = \frac{\Delta Y}{I}$	$\frac{\Delta I}{I} = \alpha\sigma$	$GC = s$	$G = \frac{\Delta Y}{Y}$
$\alpha = \frac{\Delta S}{\Delta Y}$	$\frac{\Delta I}{I} = \frac{\Delta S}{Y} \times \frac{\Delta Y}{I}$	or $\frac{\Delta Y}{Y} \times \frac{I}{\Delta Y} = \frac{S}{Y}$	$C = \frac{I}{\Delta Y}$
	$\frac{\Delta I}{I} = \frac{\Delta S}{I}$	$= \frac{I}{Y} = \frac{S}{Y}$	$s = \frac{S}{Y}$
or	or $\Delta I = \Delta S$	or $I = S$	

Given the capital-output ratio, as long as the average propensity to save is equal to the marginal propensity to save, the equality of saving and investment fulfils the conditions of equilibrium rate of growth.

Looked at from another angle, the two models are similar. Harrod's *s* is

Domar's α. Harrod's warranted rate of growth (*Gw*) is Domar's full employment rate of growth (ασ). Harrod's *Gw* = *s/Cr* ≡ Domar's ασ.

$$\text{To prove it } \alpha = \frac{S}{Y} \text{ or } S = \alpha Y \qquad \ldots(1)$$

$$\sigma = \frac{\Delta Y}{I} \text{ or } \Delta Y = I\alpha \qquad \ldots(2)$$

Since *S* = *I*, and substituting *S* for *I* in equation (2), we have

$$\Delta Y = \alpha Y \sigma \qquad [\because S = \alpha Y]$$

$$\text{or} \quad \frac{\Delta Y}{Y} = \alpha\sigma \qquad \ldots(3)$$

$$\therefore \quad Gw = \alpha\sigma \qquad \left(\text{Since } Gw = \frac{\Delta Y}{Y}\right)$$

We have proved mathematically that Harrod's *Gw* is the same as Domar's ασ . But, in reality, Domar's rate of growth α*s* is Harrod's *Gw*, and Domar's ασ is Harrod's *Gn*. In Domar's model *s* is the annual productive capacity of newly created capital which is greater than σ which is the net potential social average productivity of investment. It is the lack of labour and other factors of production which reduces Domar's growth rate from αs to ασ. Since labour is involved in σ, therefore Domar's potential growth rate resembles Horrod's natural rate. We may also say that the excess of *s* over σ in Domar's model expresses the devise of *Gw* over *Gn* in Harrod's model.

Points of Difference. There are, however, important differences in the two models.

(1) Domar assigns a key role to investment in the process of growth and emphasises on its dual character. But Harrod regards the level of income as the most important factor in the growth process. Whereas Domar forges a link between demand and supply of investment, Harrod, on the other hand, equates demand and supply of saving.

(2) The Domar model is based on one growth rate ασ. But Harrod uses three distinct rates of growth: the actual rate (*G*), the warranted rate (*Gw*) and the natural rate (*Gn*).

(3) Domar uses the reciprocal of marginal capital-output ratio, while Harrod uses the marginal capital-output ratio. In this sense Domar's $\sigma = I/Cr$ of Harrod.

(4) Domar gives expression to the multiplier but Harrod uses the accelerator about which Domar appears to say nothing.

(5) The formal identity of Harrod's *Gw* equation and Domar's equation is maintained by Domar's assumption that $\Delta I/I = \Delta Y/Y$. But Harrod does not make such assumptions. In Harrod's equilibrium equation *Gw*, there is neither any explicit or implicit reference to ΔI or I. It is, however, in his basic equation $G=s/C$ that there is an implicit reference to I, since C is defined as $I/\Delta Y$. But there is, no explicit or implicit reference to ΔI.

(6) For Harrod the business cycle is an integral part of the path of growth and for Domar it is not so but is accommodated in his model by allowing σ (average productivity of investment) to fluctuate.

(7) While Domar demonstrates the technological relationship between capital accumulation and subsequent full capacity growth in output, Harrod shows in addition a behavioural relationship between rise in demand and hence in current output on the one hand, and capital accumulation on the other. In other words, the former does not suggest any behaviour pattern for entrepreneurs and the proper change in investment comes exogenously, whereas the latter assumes a behaviour pattern for entrepreneurs that induces the proper change in investment.

LIMITATIONS OF THESE MODELS

Some of the conclusions depend on the crucial assumptions made by Harrod and Domar which make these models unrealistic:

(l) The propensity to save (α or s) and the capital-output ratio (σ) are assumed to be constant. In actuality, they are likely to change in the long run and thus modify the requirements for steady growth. A steady rate of growth can, however, be maintained without this assumption. As Domar himself writes, "This assumption is not necessary for the argument and

that the whole problem can be easily reworked with variable α and *σ.*"

(2) The assumption that labour and capital are used in fixed proportions is untenable. Generally, labour can be substituted for capital and the economy can move more smoothly towards a path of steady growth. Infact, unlike Harrod's, model, this path is not so unstable that the economy should experience chronic inflation or unemployment if *G* does not coincide with *Gw.*

(3) The two models also fail to consider changes in the general price level. Price changes always occur over time and may stabilize otherwise unstable situations. According to Meier and Baldwin, "If allowance is made for price changes and variable proportions in production, then the system may have much stronger stability than the Harrod model suggests."

(4) The assumption that there are no changes in interest rates is irrelevant to the analysis. Interest rates change and affect investment. A reduction in interest rates during periods of overproduction can make capital-intensive processes more profitable by increasing the demand for capital and thereby reduce excess supplies of goods.

(5) The Harrod-Domar models ignore the effect of government programmes on economic growth. If, for instance, the government undertakes a programmes of development, the Harrod-Domar analysis does not provide us with causal (functional) relationship.

(6) It also neglects the entrepreneurial behaviour which actually determines the warranted growth rate in the economy. This makes the concept of the warranted growth rate unrealistic.

(7) The Harrod-Domar models have been criticised for their failure to draw a distinction between capital goods and consumer goods.

(8) According to Professor Rose, the primary source of instability in Harrod's system lies in the effect of excess demand or supply on production decisions and not in the effect of growing capital shortage or redundancy on investment decisions.

Despite these limitations, “Harrod-Domar growth models are purely *laissez-faire* ones based on the assumption of fiscal neutrality and designed to indicate conditions of progressive equilibrium for an advanced economy.” They are important “because they represent a stimulating attempt to dynamize and secularise Keynes’ static short-run saving and investment theory, according to Kurihara.

EXERCISES

1. Critically examine Domar's steady growth model.
2. Bring out the essential conditions for knife-edge equilibrium growth in Harrod's growth model.
3. Point out the similarities and differences in Harrod and Domar growth models.

CHAPTER 40

THE SOLOW MODEL OF GROWTH

INTRODUCTION

Professor R.M. Solow[1] builds his model of economic growth as an alternative to the Harrod-Domar line of thought without its crucial assumption of fixed proportions in production. Solow postulates a continuous production function linking output to the inputs of capital and labour which are substitutable.

Assumptions

Solow builds his model around the following assumptions:

(1) One composite commodity is produced.

(2) Output is regarded as *net* output after making allowance for the depreciation of capital.

(3) There are constant returns to scale. In other words, the production function is homogeneous of the first degree.

1. "A Contribution to the Theory of Economic Growth", *Q.J.E.*, Vol. 70, 1956, pp.65-94

(4) The two factors of production, labour and capital, are paid according to their marginal physical productivities.

(5) Prices and wages are flexible.

(6) There is perpetual full employment of labour.

(7) There is also full employment of the available stock of capital.

(8) Labour and capital are substitutable for each other.

(9) There is neutral technical progress.

(10) The saving ratio is constant.

THE MODEL

Given these assumptions. Solow shows in his model that with variable technical coefficient there would be a tendency for capital-labour ratio to adjust itself through time in the direction of equilibrium ratio. If the initial ratio of capital to labour is more, capital and output would grow more slowly than labour force and vice versa. Solow's analysis is convergent to equilibrium path (steady state) to start with any capital-labour ratio.

Solow takes output as a whole, the only commodity, in the economy. Its annual rate of production is designated as *Y(t)* which represents the real income of the community, part of it is consumed and the rest is saved and invested. That which is saved is a constant *s*, and the rate of saving is *sY(t). K(t)* is the stock of capital. Thus net investment is the rate of increase of this stock of capital, *i.e.*, *dk/dt* or $\dot{K}$. So the basic identity is

$$\dot{K} = sY \qquad \ldots(1)$$

Since output is produced with capital and labour, technological possibilities are represented by the production function

$$Y = F(K, L) \qquad \ldots(2)$$

that shows constant returns to scale.

Inserting equation (2) in (1), we have

$$\dot{K} = sF(K, L) \qquad \ldots(3)$$

In equation (3), *L* represents total employment.

Since population is growing exogenously, the labour force increases at a constant relative rate *n*. Thus

$$L(t) = L_{oe}^{nt} \qquad \ldots(4)$$

Solow regards *n* as Harrod's natural rate of growth in the absence of technological change; and *L(t)* as the available supply of labour at time *(t)*. The right hand side of equation (4) shows the compound rate of the growth of labour force from period *0* to period *t*. Alternatively, equation (4) can be regarded as a supply curve of labour. "It says that the exponentially growing labour force is offered for employment completely inelastically. The labour supply curve is a vertical line, which shifts to the right in time as the labour force grows according to (4). Then the real wage rate adusts so that all available labour is employed, and the marginal productivity equation determines the wage rate which will actually rule."

By inserting equation (4) in (3), Solow gives basic equation

$$\dot{K} = sF(K,\ L_{oe}^{nt}) \qquad \ldots(5)$$

He regards this basic equation as determining the time path of capital accumulation, $\dot{K}$, that must be followed if all available labour is to be fully employed. It provides the time profile of the community's capital stock which will fully employ the available labour. Once the time paths of capital stock and of the labour force are known, the corresponding time path of real output can be computed from the production function.

Possible Growth Patterns. In order to find out if there is always a capital accumulation path consistent with any rate of growth of the labour force towards steady state, Professor Solow introduces his fundamental equation[2]

$$\dot{r} = sF(r,1) - nr \qquad \ldots(6)$$

In this equation *r* is the ratio of capital to labour (*K/L*), *n* is the relative rate of change of the labour force ($\dot{L}/L$). The function *sF*(*r*,1) represents

output per worker as a function of capital per worker. In other words,it is the total product curve as varying amounts r of capital are employed with one unit of labour. The equation (6) itself states that the rate of change of capital-labour ratio ($\dot{r}$) is the difference of two terms, one representing the increment of capital [$sF(r,1)$] and the other increment of labour (nr).

Solow illustrates diagrammatically possible growth patterns based on his fundamental equation (6).

In Fig. 1, the ray through the origin is the function nr. The other curve represents the function $sF(r,1)$. It is so drawn as to show diminishing marginal productivity of capital. At the point of intersection of the two curves $nr=sF(r, 1)$, and $\dot{r}=O$. Then $r=r'$. When $r=O$, the capital-labour ratio is a constant and the capital stock must expand at the same rate as the labour force, *i.e.*, n. Once the capital-labour ratio r' is established, it will be maintained, and capital and labour will grow in proportion. Assuming constant returns to scale, real output will also grow at the same relative rate n, and output per head of labour force will be constant. At r' there will be the balanced growth equilibrium.

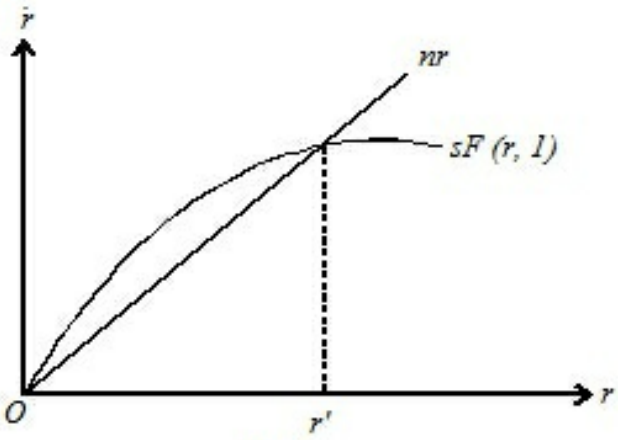


Fig. 1

---

2. We arrive at the fundamental equation thus:

Equation (5) helps to study the behaviour of the capital-labour ratio (K/L). To do this, he introduces a new variable r for the capital-labour ratio. Thus $r=K/L$ or $K=rL$. Substituting (4) into this expression, we have

$$K=r\,L_{0e}^{nt}$$

Differentiating with respect to time to get an equation for the rate of change of capital stock, the product rule of derivates gives us

$$\frac{dK}{dt} = L_{0e}^{nt}\frac{d}{dt}\,r + r\,\frac{d}{dt}L_{0e}^{nt}$$

$$\dot{K} = \dot{r}L_{oe}^{nt} + nrL_{oe}^{nt}$$

$$= (\dot{r} + nr)L_{oe}^{nt} \qquad \text{(By taking } L_{oe}^{nt} \text{ as common)}$$

Substituting (5) into it, we have

$$(\dot{r} + nr)\ L_{oe}^{nt} = sF(K,\ L_{oe}^{nt})$$

This tells us how capital is growing assuming that labour is fully employed and a fraction, *s*, of full employment output is saved in each period.

Assuming constant returns to scale which means that the production function is homogeneous of degree 1 we divide the above relation by L_{oe}^{nt} to get

$$(\dot{r} + nr) = sF\left(\frac{K}{L_{oe}^{nt}}, 1\right)$$

Taking *nr* to R.H.S.,

$$\dot{r} = sF\left(\frac{K}{L_{oe}^{nt}}, 1\right) - nr$$

Finally, writing *r* for the capital-labour ratio $\frac{K}{L_{oe}^{nt}}$ we get the Solow fundamental equation (6)

$$\dot{r} = sF(r,1) - nr$$

What will be the behaviour of the capital-labour ratio if there is a divergence between r' and *r*. If *r* lies to the right of r' or $r > r'$ then $nr > sF(r,1)$, and *r* will decrease toward r'. On the contrary, if *r* lies to the left of r' or $r < r'$, $nr < sF(r,1)$, and *r* will increase toward r'. Thus the equilibrium value r' is stable. "Whatever the initial value of the capital-labour ratio, the system will develop *toward* a state of balanced growth at the natural rate... If the initial capital stock is below the equilibrium ratio, capital and output will grow at a faster pace than the labour force until the equilibrium ratio is approached. If the initial ratio is above the equilibrium value, capital and output will grow more slowly than the labour force. The growth of output is always intermediate between those of labour and capital."

But the strong stability shown in the above figure is not inevitable. It depends on the shape of the productivity curve $sF(r,1)$. In Fig. 2 the

productivity curve $sF(r,1)$ intersects the ray curve nr at three points r_1, r_2 and r_3. But r_1 and r_3 are stable equilibrium positions because the total productivity curve $sF(r, 1)$ is above nr but at r_2 it is below nr. Therefore, r_2 is an unstable equilibrium position. "Depending on the initial observed capital-labour ratio, the system will develop either to balanced growth at capital-labour ratio r_1 or r_3. In either case labour supply, capital stock and real output will asymptomatically expand at rate n, but around r_1 there is less capital than around r_3, hence the level of output per head will be lower in the former case than in the latter. The relevant balanced growth equilibrium is at r_1 for an initial ratio anywhere between O and r_2, it is at r_3 for any initial ratio greater than r_2. The ratio r_2 is itself an equilibrium growth ratio, but an unstable one, any accidental disturbance will be magnified over time. Figure 2 has been drawn so that production is possible without capital; hence the origin is not an equilibrium 'growth' configuration."

Solow points out that Fig. 2 does not exhaust all possibilities. He shows two more possibilities, as shown in Fig. 3. The ray nr depicts the equilibrium growth path where the warranted and natural rates of growth are equal. The curve $s_1F'(r,1)$ which is above nr represents a highly productive system in which capital and income increase more rapidly than the labour supply. In this system, which is of perpetual full employment, income and saving increase so much that the capital-labour ratio increases limitlessly. On the other hand, the curve $S_2F''(r,1)$ depicts a highly unproductive system in which the full employment path leads to ever diminishing per capita income. However, aggregate income rises in his system because net investment is always positive and the labour supply is

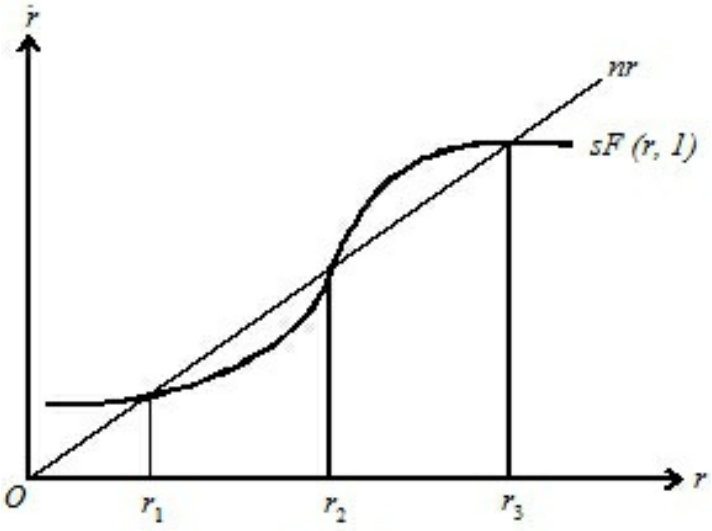


Fig. 2

increasing. It is to be noted that both the systems have diminishing marginal productivity throughout.

Professor Solow concludes his model thus: "When production takes place under the usual neo-classical conditions of variable proportions and constant returns to scale, no simple opposition between natural and warranted rates of growth is possible. There may not be...any knife-edge. The system can adjust to any given rate of growth of the labour force, and eventually approach a state of steady proportional expansion," *i.e.*,

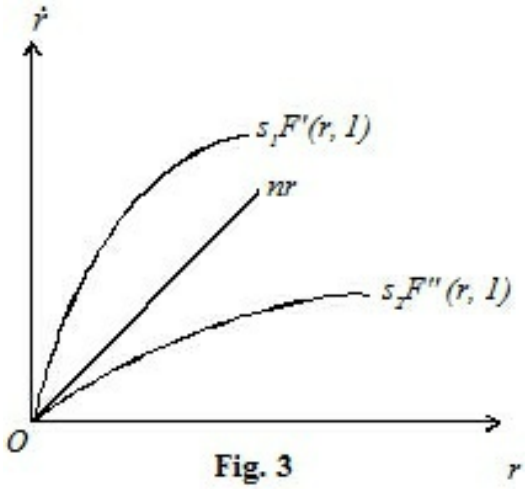


Fig. 3

$$\frac{\Delta K}{K} = \frac{\Delta L}{L} = \frac{\Delta Y}{Y}$$

A Critical Appraisal

The Solow model is a major improvement over the Harrod-Domar model. The Harrod-Domar model is at best a knife-edge balance in a long-run economic system where the saving ratio, the capital-output ratio, and the rate of increase of the labour force are the key parameters. If the magnitudes of these parameters were to slip even slightly from the dead centre, the consequences would be either growing unemployment or chronic inflation. In Harrod's terminology, this balance is poised on the equality of *Gw* (which depends on the saving and investing habits of households and firms) and *Gn* (which depends, in the absence of technical change, on the increase of the labour force). According to Solow, this delicate balance between *Gw* and *Gn* flows from the crucial assumption of fixed proportions in production whereby there is no possibility of substituting labour for capital. If this assumption is abandoned, the knife-edge balance between *Gw* and *Gn* also disappears with it. He, therefore, builds a model of long-run growth without the assumption of fixed proportions in production demonstrating steady state growth.

Solow is a pioneer in constructing the basic neo-classical model where he retains the main features of the Harrod-Domar model like homogeneous capital, proportional saving function and a given growth rate in the labour force. He takes a continuous production function, which has come to be known as the neo-classical production function, in analysing the process of growth. The assumption of substitutability between labour and capital gives the growth process an adjustability and provides a touch of realism. Unlike the Harrod-Domar model, he demonstrates steady-state growth paths. Last but not the least, the long-run rate of growth is determined by an expanding labour force and technical progress. Thus Professor Solow has successfully shunted aside all the difficulties and rigidities which go into the modern Keynesian income analysis.

Weaknesses

His "purpose was to examine what might be called the tight-rope view of economic growth and to see where more flexible assumptions about production would lead a simple model." Despite this assertion by Solow, his model is weak in many respects, according to Prof. Amartya Sen.

1. The Solow model takes up only the problem of balance between Harrod's *Gw* and *Gn* and leaves out the problem of balance between *G* and *Gw*.

2. There is the absence of an investment function in Solow's model and once it is introduced, the Harrodian problem of instability quickly reappears by the Solow model. Thus, according to Sen, the assumption of substitutability between labour and capital does not seem to be a key difference between neo-classical and neo-Keynesian studies of growth, and the main difference seems to lie in the investment function and the consequent failure to assign a major role to entrepreneurial expectations about the future.

3. The Solow model is based on the assumption of labour-augmenting technical progress. It is, however, a special case of Harrod-neutral technical progress of the Cobb-Douglas production function type which does not possess any empirical justification.

4. Solow assumed flexibility of factor prices which may bring difficulties in the path towards steady growth. For instance, the rate of interest may be prevented from falling below a certain minimum level due to te problem of liquidity trap. This may, in turn, prevent the capital-output ratio from rising to a level necessary for attaining the path of equilibrium growth.

5. The Solow model is based on the unrealistic assumption of homogeneous and malleable capital. As a matter of fact, capital goods are highly heterogeneous and thus pose the problem of aggregation. Consequently, it is not easy to arrive at the steady growth path when there are varieties of capital goods.

6. Solow leaves out the causative of technical progress and treats the latter as an exogenous factor in the growth process. He thus ignores the problems of inducing technical progress through the process of learning, investment in research, and capital accumulation.

EXERCISES

1. Explain critically Solow's Model of Growth.
2. Explain Solow's model of Economic Growth. In what ways it is an improvement over the Harriod-Domar model ?

CHAPTER 41

THE SOLOW-SWAN MODELS OF ECONOMIC GROWTH

THE SOLOW-SWAN MODEL

The Solow-Swan model[1] of economic growth postulates a continuous production function linking output to the inputs of capital and labour which leads to the steady state equilibrium of the economy.

Its Assumptions

It is based on the following assumptions :

1. One composite commodity is produced.

2. Output is regarded *net* output after making allowance for the depreciation of capital.

3. There are constant returns to scale.

4. There are diminishing returns to an individual input.

1. R.M. Solow, "A Conribution to the Theory of Economic Growth, *"Q.J.E.* Feb. 1956. T.W. Swan, "Economic Growth and Capital Accumulation," *Economic Record,* Nov. 1956.

5. The two factors of production, labour and capital, are paid according to their marginal physical productivities.

6. Prices and wages are flexible.

7. There is perpetual full employment of labour.

8. There is also full employment of the available stock of capital.

9. Labour and capital are substitutable for each other.

10. There is no technical progress.

11. The saving ratio is constant.

12. Saving equals investment.

13. Capital depreciates at the constant rate, *d*.

14. Population grows at a constant rate, *n*.

The Model

Given these asumptions, with unchanging technical progress, the production function is

$$Y = F\,(K,L)$$

where Y is income or output, K is capital and L is labour. The condition of constant returns to scale implies that if we divide by L, the production function can be written as

$$Y / L = F\,(K/L,1) = L.\ f(k)$$

where $Y = Y/L$ is output or income per worker, $k = K/L$ is the capital-labour ratio, and the function $f(k) = f(k, 1)$. Thus the production function can be expressed as

$$y = f(k) \qquad \ldots(2)$$

In the Solow-Swan model, saving is a constant fraction, *s*, of income. So saving per worker is *sy*. Since income equals output,

$$sy = sf(k) \quad \ldots(3)$$

The investment required to maintain capital per worker *k,* depends on population growth, and the depreciation rate, *d.* Since it is assumed that population grows at a constant rate *n,* the capital stock grows at the rate *n.k* to provide capital to the growing population. Since depreciation is a constant, *d,* per cent of the capital stock, *d. k* is the investment needed to replace worn-out capital. This depreciation investment per worker *d.k* is added to *nk,* the investment per worker to maintain capital-labour ratio for the growing population,

$$(nk + dk) = (n + d)k \quad \ldots(4)$$

which is the investment required to maintain capital per worker.

The net change in capital per worker (capita-labour ratio) $\dot{k}$ over time is the excess of saving per worker over the required investment to maintain capital per worker,

$$\dot{k} = sf(k) - (n + d)k \quad \ldots(5)$$

This is the fundamental equation for the Solow-Swan model, where the steady state corresponds to $\dot{k} = 0$. The economy reaches a steady state when

$$sf(k) = (n + d)\, k \quad \ldots(6)$$

The Solow-Swan model is explained in Fig. 1.

Output per worker *y* is measured along the vertical axis and capital per worker (capital-labour ratio), k, is measured along the horizontal axis. The $y = f(k)$ curve is the production function which shows that output per worker increases at a diminishing rate as *k* increases due to the law of diminishing returns. The $sf(k)$ curve represents saving per worker. The $(n + d)k$ is the investment requirement line from the origin with a positive slope equal to $(n+d)$ The steady state level of capital, $\dot{k}$, is determined where the $sf(k)$ curve intersects the $(n+d)k$ line at point E. The steady state income is y with output per worker $\dot{k}P$, as measured by point *P* on the

production function $y = f(k)$.

In order to understand why $\dot{k}$ is a steady state situation, suppose the economy starts at the capital-labour ratio k_1. Here saving per worker k_1B exceeds the investment required to keep the capital-labour ratio constant, k_1A, ($k_1B > k_1A$). Thus, k and y increase until k is reached when the economy is in the steady state at point E. Alternatively, if the capital-labour ratio is k_2, the saving per worker, k_2C, will be less than the investment required to keep the capital-labour ratios constant, k_2D, ($k_2C < k_2D$). Thus y will fall as k falls to k and the economy reaches the steady state E.

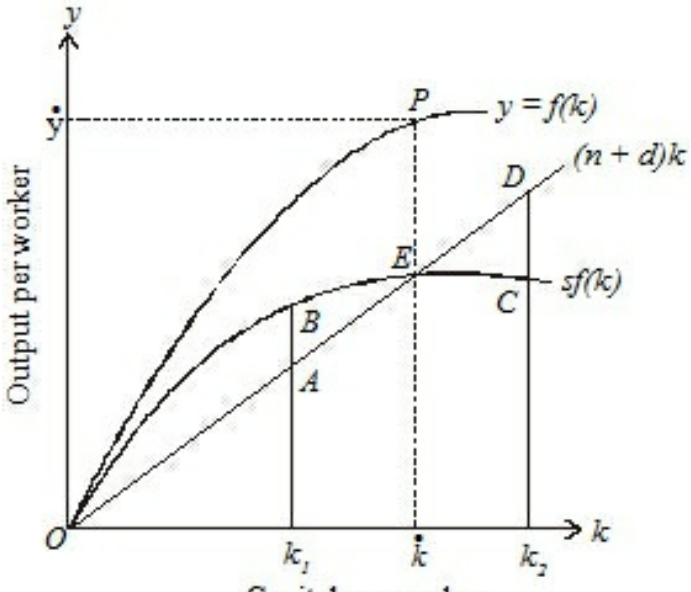


Fig. 1

The Solow-Swan model shows that the growth process is stable. No matter where the economy starts, forces exist that will push the economy over time to a steady state.

Growth with Saving

An important conclusion of the Solow-Swan model is that the growth rate does not depend upon the saving rate. In the steady state, both k and y being constant, the growth rate is not affected by the saving rate. This is explained in Fig. 2 where $\dot{k}$, is the steady state capital per worker and $\dot{y}$ is output per worker when the $sf(k)$ curve intersects the $(n+d)$ $\dot{k}$ curve at point E. An increase in the saving rate from s to s_1 shifts the saving curve $sf(k)$ upward to s_1 $f(k)$. The new steady state point is E_1.

When the saving rate increases forms s to s_1 with no change in the growth rate of labour force (n), the capital per worker will continue to rise to $\dot{k}_1$

which will raise output per worker to $\dot{y}_1$ and so will the growth rate of output increase. But this process continues at a diminishing rate in the transition period. As a result, the initial growth rate of output is restored over the long run at the new steady state equilibrium point E_1 where $(n+d)k = s_1 f(k)$.

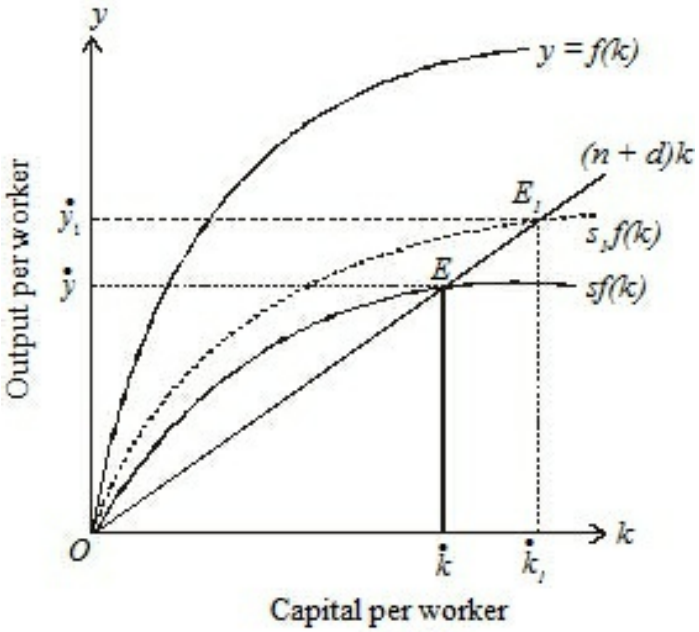


Fig. 2

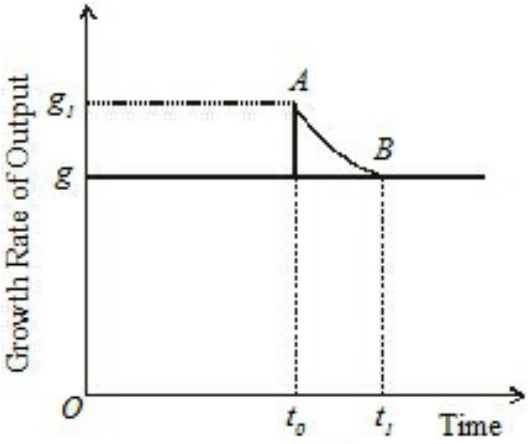


Fig. 3

After this point, there will be no further increase in output per worker because the growth rate of labour force (n) does not change and the long-run growth rate of output also remains at the same level. Figure 3 depicts the effect on the growth rate of output when there is increase in the saving rate. The saving rate increases at time t_0. Initially, the growth rate of output rises from g to g_1. This is the transition period in which output per worker is increasing from $\dot{y}$ to $\dot{y}_1$ and capital per worker from $\dot{k}$ to $\dot{k}_1$, as shown in Fig. 2 But at time t_1, the initial equilibrium growth rate is restored with the fall in the growth rate of output from point *A* to *B.*

IMPLICATIONS OF THE MODEL

There are some important implications or predictions of the Solow-Swan

model of growth.

1. The growth rate of output in steady state is exogenous and is independent of the saving rate and technical progress.

2. If the saving rate increases, it increases the output per worker by increasing the capital per worker, but the growth rate of output is not affected.

3. Another implication of the model is that growth in per capita income can either be achieved by increased saving or reduced rate of population growth. This will hold if depreciation is allowed in the model.

4. Another prediction of the model is that in the absence of continuing improvements in technology, growth per worker must ultimately cease. This prediction follows from the assumption of diminishing returns to capital.

5. This model predicts conditional convergence. All countries having similar characteristics like saving rate, population growth rate, technology, etc. that affect growth will converge to the same steady state level. It means that poor countries having the same saving rate and level of technology of the rich countries will reach the same steady state growth rates in the long run.

EXERCISES

1. Explain the Solow-Swan model of economic growth. What are its implications?
2. Explain the Solow-Swan model of economic growth with and without saving.

CHAPTER 42

THE ENDOGENOUS GROWTH THEORY

INTRODUCTION

The endogenous growth theory was developed as a reaction to omissions and deficiencies in the Solow-Swan neoclassical growth model. It is a new theory which explains the long-run growth rate of an economy on the basis of endogenous factors as against exogenous factors of the neoclassical growth theory. The Solow-Swan neoclassical growth model explains the long-run growth rate of output based on two exogenous variables: the rate of population growth and the rate of technological progress and that is independent of the saving rate. As the long-run growth rate depended on exogenous factors, the neoclassical theory had few policy implications. As pointed out by Romer, "In models with exogenous technical change and exogenous population growth, it never really mattered what the government did." The new growth theory does not simply criticise the neoclassical growth theory. Rather, it extends the latter by introducing endogenous technical progress in growth models. The endogenous growth models have been developed by Arrow[1], Romer[2] and Lucas[3], among other economists. We briefly study their main features, criticisms and policy implications.

THE ENDOGENOUS GROWTH MODELS

The endogenous growth models emphasise technical progress resulting from the rate of investment, the size of the capital stock, and the stock of human capital.

Assumptions

The new growth theories are based on the following assumptions:

1. There are many firms in a market.

2. Knowledge or technological advance is a non-rival good.

3. There are increasing returns to scale to all factors taken togather and constant returns to a single factor, at least for one.

4. Technological advance comes from things people do. This means that technological advance is based on the creation of new ideas.

5. Many individuals and firms have market power and earn profits from their discoveries. This assumption arises from increasing returns to scale in production that leads to imperfect competition.

As a matter of fact, these are the requirements of an endogenous growth theory.

Given these assumptions, we explain the three main models of endogenous growth.

1. Arrow's Learning by Doing and Other Models

The Arrow Model. Arrow was the first economist to introduce the concept of learning by doing in 1962 by regarding it as endogenous in the growth process. His hypothesis was that at any moment of time new capital goods incorporate all the knowledge then available based on accumulated experience, but once built, their productive deficiencies cannot be changed by subsequent learning. Arrow's model in a simplified form can be written as

$$Y_i = A(K)F(K_i, L_i)$$

where Y_i denotes output of firm i, K_i donates its stock of capital, L_i denotes its stock of labour, K without a subscript denotes the aggregated

stock of capital and *A* is the technology factor. He showed that if the stock of labour is held constant, growth ultimately comes to a halt because socially very little is invested and produced. Therefore, Arrow did not explain that his model could lead to sustained endogenous growth.

The Levhari-Sheshinski Model. Arrow's model has been generalised and extended by Levhari and Sheshinski. They emphasise the spillover effects of increased knowledge as the source of knowledge. They assume that the source of knowledge or learning by doing is each firm's investment. An increase in a firm's investment leads to a parallel increase in its level of knowledge. An other assumption is that the knowledge of a firm is a public good which other firms can have at zero cost. Thus knowledge has a non-rival character which spillsover across all the firms in the economy. This stems from the fact that each firm operates under constant returns to scale and the economy as a whole is operating under increasing returns to scale.

---

1. Kenneth J. Arrow, "The Economic Implications of Leaning by Doing," *R.E.S.*, June 1962.

2. Paul M. Romer, "Increasing Returns and Long-run Growth," *J.P.E.*, October 1986; Endogenous Technological Change," *J.P.E.*, October 1990.

3. Robert E. Lucas, Jr, "On the Mechanics of Economic Development," *Journal of Monetary Economics*," July 1988.

In the Levhari-Sheshinski Model, endogenous technical progress in terms of knowledge or learning by doing is reflected in an upward raising of the production function and economic growth is explained "in the context of aggregate increasing returns being consistent with competitive equilibrium."

The King-Robson Model. King and Robson emphasise *learning by watching* in their technical progress function. Investment by a firm represents innovation to solve the problems it faces. If it is successful, the other firms will adapt the innovation to their own needs. Thus externalities resulting from learning by watching are a key to economic growth. The King and Robson study shows that innovation in one sector of the economy has the contagion or demonstration effect on the productivity of other sectors, thereby leading to economic growth. They conclude that

multiple steady state growth paths exist, even for economies having similar initial endowments, and policies that increase investment should be pursued.

The Romer Model. Romer in his first paper on endogenous growth in 1986 presented a variant on Arrow's model which is known as *learning by investment*. He assumes creation of knowledge as a side product of investment. He takes knowledge as an input in the production function of the following form

$$Y = A(R)\, F\,(R_i\, ,K_i\, ,L_i\,)$$

where Y is aggregate output; A is the public stock of knowledge from research and development R; R_i is the stock of results from expenditure on research and development by firm i; and K_i and L_i are capital stock and labour stock of firm i respectively.

He assumes the function F homogeneous of degree one in all its inputs R_i, K_i, and L_i, and treats R_i as a rival good.

Romer took three key elements in his model, namely externalities, increasing returns in the production of output and diminishing returns in the production of *new knowledge*. According to Romer, it is spillovers from research efforts by a firm that leads to the creation of new knowledge by other firms. In other words, new research technology by a firm spillsover instantly across the entire economy.

In his model, new knowledge is the ultimate determinant of long-run growth which is determined by investment in research technology. Research technology exhibits diminishing returns which means that investments in research technology will not double knowledge. Moreover, the firm investing in research technology will not be the exclusive beneficiary of the increase in knowledge. The other firms also make use of the new knowledge due to the inadequacy of patent protection and increase their production. Thus the production of goods from increased knowledge displays increasing returns and competitive equilibrium is consistent with increasing aggregate returns owing to externalities. Thus

Romer takes investment in research technology as endogenous factor in terms of the acquisition of new knowledge by rational profit maximisation firms.

2. The Lucas Model

Uzawa developed an endogenous growth model based on investment in human capital which was used by Lucas. Lucas assumes that investment on education leads to the production of human capital which is the crucial determinant in the growth process. He makes a distinction between the *internal effects* of human capital where the individual worker undergoing training becomes more productive, and *external effects* which spillover and increase the productivity of capital and of other workers in the economy. It is investment in human capital rather than physical capital that has spillover effects that increase the level of technology. Thus the output for firm *i* takes the form

$$Y_i = A(K_i).(H_i).H^e$$

where A is the technical coefficient, K_i and H_i are the inputs of physical and human capital used by firms to produce goods Y_i . The variable H is the economy's average level of human capital. The parameter e represents the strength of the external effects from human capital to each firm's productivity.

In the Lucas model, each firm faces constant returns to scale, while there are increasing returns for the whole economy. Further, learning by doing or on-the-job training and spillover effects involve human capital. Each firm benefits from the *average* level of human capital in the economy, rather than from the *aggregate* of human capital. Thus it is not the accumulated knowledge or experience of other firms but the average level of skills and knowledge in the economy that are crucial for economic growth. In the model, technology is endogenously provided as a side effect of investment decisions by firms. Technology is treated as a public good from the point of view of its users. As a result, firms can be treated as price takers and there can be an equilibrium with many firms as under perfect competition.

3. Romer's Model of Technological Change

Romer's model of *Endogenous Technical Change* of 1990 identifies a research sector specialising in the production of ideas. This sector invokes human capital alongwith the existing stock of knowledge to produce ideas or new knowledge. To Romer, ideas are more important than natural resources. He cites the example of Japan which has very few natural resources but it was open to new western ideas and technology. It imported machines from the United States during the Meija era, dismantled them to see how they worked and manufactured their better prototypes. Therefore, ideas are essential for the growth of an economy. These ideas relate to improved designs for the production of producer durable goods for final production.

In the Romer model, new knowledge enters into the production process in three ways. *First*, a new design is used in the intermediate goods sector for the production of a new intermediate input. *Second*, in the final sector, labour, human capital and available producer durables produce the final product. *Third*, and a new design increases the total stock of knowledge which increases the productivity of human capital employed in the research sector.

Its Assumptions

The Romer model is based on the following assumptions:

1. Economic growth comes from technological change.

2. Technological change is endogenous.

3. Market incentives play an important role in making technological changes available to the economy.

4. Invention of a new design requires a specified amount of human capital.

5. The aggregate supply of human capital is fixed.

6. Knowledge or a new design is assumed to be *partially excludable and*

retainable by the firm which invented the new design. It means that if an inventor has a patented design for a machine, no one can make or sell it without the agreement of the inventor. On the other hand, other inventors are free to spend time to study the patented design for the machine and acquire knowledge that helps in the design of such a machine. Thus patents provide incentives to firms to engage in research and development, and other firms can also benefit from such knowledge. When there is partial excludability, investment in research and development leading to an invention by a firm can only bring in quasi-rent.

7. Technology is a non-rival input. Its use by one firm does not prevent its use by another.

8. The new design can be used by firms and in different periods without additional costs and without reducing the value of the input.

9. It is also assumed that the low cost of using an existing design reduces the cost of creating new designs.

10. When firms make investments on research and development and invent a new design, there are externalities that are internalised by private agreements.

The Model

Given these assumptions, the Romer model can be explained in terms of the following techknological production function,

$$\Delta A = F\,(K_A, H_A, A)$$

where ΔA is the increasing technology, K_A is the amount of capital invested in producing the new design (or technology), H_A is the amount of human capital (labour) employed in research and development of the new design, A is the existing technology of designs, and F is the production function for technology.

The production function shows that technology is endogenous when more human capital is employed for research and development of new designs,

then technology increases by a larger amount, *i.e.*, A is greater. If more capital is invested in research laboratories and equipment to invent the new design, then technology also increases by a larger amount *i.e.*, ΔA is more. Further, the existing technology, A, also leads to the production of new technology, ΔA. Since it is assumed that technology is a non-rival input and partially excludable, there are positive spillover effects of technology which can be used by other firms. Thus the production of new technology (knowledge or idea) can be increased through the use of physical capital, human capital and existing technology.

Criticisms of Endogenous Growth Theory

Despite the fact that the new growth theory has been regarded as an improvement over the new classical growth theory, still it has many critics.

1. According to Scott and Auerbach, the main ideas of the new growth theory can be traced to Adam Smith and increasing returns to Marx's analysis.

2. Srinivasan does not find any thing new in the new growth theory because increasing returns and endogeneity of variables have been taken from the neoclassical and Kaldor's models.

3. Fisher criticises the new growth theory for depending only on the production function and the steady state.

4. To Olson, the new growth theory lays too much emphasis on the role of human capital and neglects the role of institutions.

5. In the various models of new growth theory, the difference between physical capital and human capital is not clear. For instance, in Romer's model, capital goods are the key to economic growth. He assumes that human capital accumulates and when it is embodied in physical capital then it becomes a driving force. But he does not clarify which is the driving force.

6. By using secondary school enrollment as a proxy for human capital in

their model, Mankiw, Romer and Weil find that physical and humann capital accumulation cannot lead to perpetual economic growth.

POLICY IMPLICATIONS OF ENDOGENOUS GROWTH THEORY

The endogenous growth theory has important policy implications for both developed and developing economies.

1. This theory suggests that convergence of growth rates per capita of developing and developed countries can no longer be expected to occur. The increasing returns to both physical and human capital imply that the rate of return to investment will not fall in developed countries relative to developing countries. In fact, the rate of return to capital in developed countries is likely to be higher than that in developing countries. Therefore, capital need not flow from the developed to the developing countries and actually the reverse may happen.

2. Another implication is that the measured contribution of both physical and human capital to growth may be larger than suggested by the Solow residual model. Investment on education or research and development of a firm has not only a positive effect on the firm itself but also spillover effects on other firms and hence on the economy as a whole. This suggests that the residual attributed to technical change in the Solow growth accounting may be actually much smaller.

3. One of the important implications is that it is not necessary that economies having increasing returns to scale must reach a steady state level of income growth, as suggested by the Solow-Swan model. When there are large positive externalities from new investment on research and development, it is not necessary for diminishing returns to start. So the growth rate of income does not slow down and the economy does not reach steady state. But an increase in the saving rate can lead to a permanent increase in the growth rate of the economy.

4. This further implies that countries having greater stocks of human capital and investing more on research and development will enjoy a faster rate of economic growth. This may be one of the reasons for the

slow growth rate of certain developing countries.

EXERCISES

1. Explain the main features of the endogenous growth theory. What are its policy implications?
2. Explain Arrow's Learning by Doing and Romer's Learning by Investment models.
3. Explain critically Romer's Model of Technological change.

CHAPTER 43

STEADY STATE GROWTH

MEANING

The concept of steady state growth is the counterpart of long-run equilibrium in static theory. It is consistent with the concept of *equilibrium growth*. In steady state growth all variables, such as output, population, capital stock, saving, investment, and technical progress, either grow at constant exponential rate, or are constant.

Taking different variables, some of the neo-classical economists have given their interpretations to the concept of steady state growth. To begin with Harrod, an economy is in a state of steady growth when $Gw=Gn$. Joan Robinson described the conditions of steady state growth as Golden Age of accumulation thus indicating a "mythical state of affairs not likely to obtain in any actual economy." But it is a situation of stationary equilibrium. According to Meade, in a state of steady growth, the growth rate of total income and the growth rate of income per head are constant with population growing at a constant proportionate rate, with no change in the rate of technical progress. Solow in his model demonstrates steady growth paths as determined by an expanding labour force and technical progress.

PROPERTIES OF STEADY STATE GROWTH*

The neo-classical theory of economic growth is concerned with analysing the properties of steady state growth based on the following basic *assumptions* of the Harrod-Domar model:

1. There is only one composite commodity which can be consumed or used as an input in production or can be accumulated as a capital stock.

2. Labour force grows at a constant proportional rate n.

3. Full employment prevails at all times.

4. Capital-output ratio (v) is also given.

5. Saving-income ratio (s) is constant.

6. There are fixed coefficients of productions. In other words, there is no possibility of the substitution of capital and labour.

7. There is no technical change (m).

The neo-classical growth models discuss the properties of steady state growth by incorporating and relaxing these assumptions.

In order to discuss the properties of steady state growth, we first study the Harrod-Domar model briefly.

The Harrod-Domar model is not a steady state growth model where Gw ($= s/v$) = Gn ($=n + m$). It is one of knife-edge balance between cumulative inflation and cumulative deflation. It is only when the warranted growth rate s/v equals the natural rate of growth $n+m$, that there will be steady state growth. But, s,v.n and m being independent constants, there is no valid reason for the economy to grow at full employment steady state. So we discuss the roles assigned to them one by one in neo-classical growth theory.

1. Flexibility of *n*

Economists like Joan Robinson and Kahn have shown that the presence of unemployment is compatible with steady growth. So the assumption of the growth rate of labour force at full employment is dropped. Instead, it is replaced by the condition that the growth rate of employment should not be greater than n. For steady growth it is not necessary that $s/v=n$. Rather,

equilibrium growth is compatible with $s/v<n$. This is what Kahn calls a *bastard golden age* as against Joan Robinson's golden age where $s/v=n$. In a bastard golden age, the rate of capital accumulation (s/v) is less than the growth rate of population (n), so that unemployment increases. In this age, capital stock is not growing faster because of inflationary pressures. Rising prices mean a lower real wage rate. When the real wage rate is at the tolerably minimum level, it sets a limit to the rate of capital accumulation.

2. Flexible Capital-Output Raio (v)

Now we turn to the second assumption of the Harrod-Domar model, that of a constant capital-output ratio (v). Solow and Swan have built models of steady state growth with a variable capital-output ratio.

Theoretically, the Harrod-Domar assumption of an unchanging capital-output ratio implies that the amount of capital and labour required to produce a unit of output are fixed. The neo-classical economists postulate a continuous production function linking output to the inputs of capital and labour. The other assumptions of constant returns to scale, no technical progress and constant saving ratio are retained.

---

* The following sections also relate to the *Basic Neo Classical Growth Model.*

Solow-Swan show that because of the substitutability of capital and labour and by increasing the capital-labour ratio, the capital-output ratio can be increased and hence the warranted rate s/v can be made equal to the natural rate, $n+m$. If the warranted growth rate exceeds the natural growth rate, the economy tries to break through the full employment barrier, thereby making labour more expensive in relation to capital, and making inducements to shift to labour-saving techniques. This raises the capital-output ratio and the value of s/v is reduced until it coincides with $n+m$. If, on the other hand, the warranted growht rate is less than the natural growth rate, there will be surplus labour which lowers the real wage rate in relation to the real interest rate. Consequently, more labour-intensive techniques are chosen which reduce the capital-output ratio (v) thereby raising s/v. This process continues till s/v equals $n+m$. Thus, it is the

capital-output ratio which maintains the steady state growth single-handed while *s*, *n* and *m* remain constant.

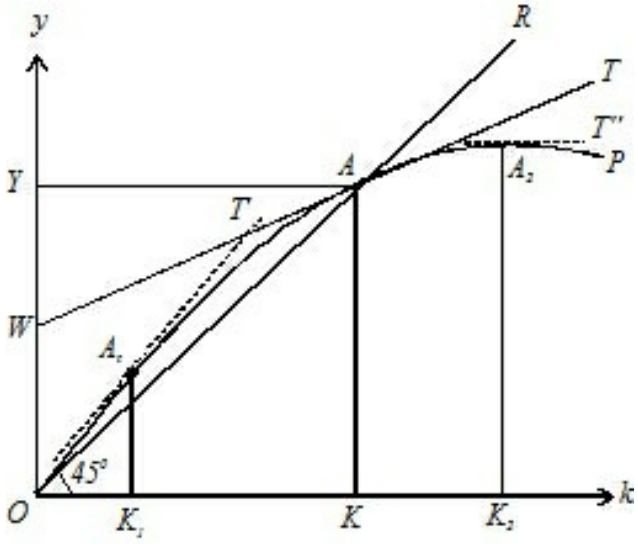


Fig. 1

This situation is explained in Fig. 1 where capital-labour ratio (or capital per man) *k*, is taken on the horizontal axis and output per man, *y*, is taken on the verticle axis. The 45° line *OR* represents capital-output ratio where the warranted growth rate equals the natural growth rate. Every point on *OR* also shows a constant capital-labour ratio. *OP* is the production function which measures the marginal productivity of capital. It also expresses the relation between output per man (*y*) and capital per man (*k*). The tangent *WT* to the production function *OP* indicates the rate of profit at point *A* corresponding to the marginal productivity of capital. It is at this point *A* that the warranted growth rate equals the natural growth rate, *i.e.*, *s*/*v*=*n*+*m*. Here the share of profit is *WY* in national, income is *OY*, and *OW* is the wage per man. Assume a situation K_2 where the stock of capital is above the equilibrium stock. It indicates that the capital-labour ratio is above the full employment equilibrium level ratio at A_2. Thus, there is some idle capital which cannot be utilised and the rate of profit declines (which can be shown by joining tangent *T''* at A_2 to the *Y*-axis where it shall be above *OW*) till it reaches point *A* of steady state growth. The opposite is the case at K_1 where the growth rate of capital accumulation is higher than that of labour force. The rate of profit increases at A_1 (which can be shown by joining the tanget *T'* to the *Y*-axis where it shall be below *OW*) till the steady state growth point *A* is reached. In the Harrod-Domar model there is a single point of equilibrium *A* on the production function *OP* because the capital-output ration (*v*) is fixed. But in the new-classical model there is a continuous production function along which the capital-output ratio is a variable and if the economy is thrown off the steady state level *A*, it will

itself return to it by variations in the capital-labour ratio. Thus the equilibrium value of K is stable.

3. Flexibility of Saving Ratio (s)

The Harrod-Domar model is also based on the assumption of a constant saving-income ratio (s). Kaldor and Pasinetti have developed the hypothesis which treats the saving-income ratio as a variable in the growth process. It is based on the classical saving function which implies that savings equal the ratio of profits to national income.

The hypothesis is that the economy consists of only two classes, the wage-earners and the profit-earners. Their savings are a function of their incomes. But the propensity to save of profit-earners(sp) is higher than that of wages-earners (sw). As a result, the overall saving ratio of the community depends on the distribution of income.

A special case of this hypothesis is where the propensity to save out of wages is zero (sw=0) and the propensity to save out of profits is positive and constant. Thus the overall propensity to save (s) is equal to the propensity to save of profit-earners (sp) multiplied by the ratio of profits (π) to the national income (Y), *i.e.*, $S = sp.\pi/Y$. This is th classical saving function. There is also the 'extreme' classical saving function where all wages are consumed (sw=0) and all profits are saved (sp=1). Hence the saving-income ratio $s = \pi/Y$.

With a constant capital-output ratio (v) and a variable saving-income ratio (s), steady state growth can be maintained through the distribution of income. So long as the saving-income ratio (s) required to satisfy the condition $s/v=n+m$ is not less than the propensity to save of wage-earner (sw=o) and not greater than the propensity to save of profit-earners (sp=1), steady state growth will be maintained.

4. Flexible Saving Ratio (s) and Flexible Capital-Output Ratio (v)

Steady state growth can also be shown by taking both the saving-income ratio and the capital-output ratio as variables. With the classical saving function given by $sp.\pi/Y$, the warranted growth rate s/v can be written as :

$$Gw = \frac{s}{v} = \frac{sp\pi}{Y} \cdot \frac{Y}{K} = \frac{sp\pi}{K} \qquad \left[\begin{array}{c} \because s = sp.\pi / Y \\ v = K / Y \end{array}\right]$$

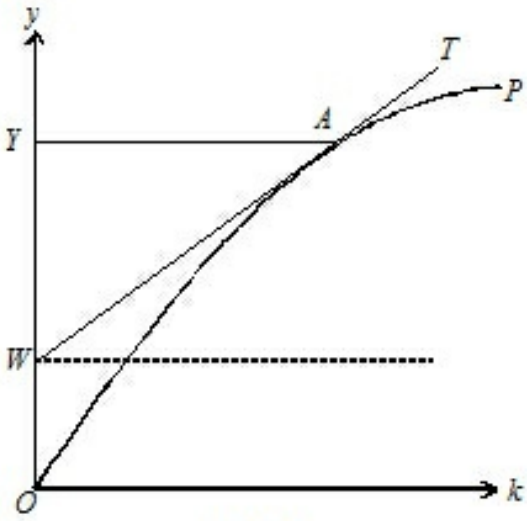


Fig. 2

where π/K is the rate of profit on captial which can be denoted by *r*. Thus the warranted rate becomes *spr*. For steady state growth, *spr* = *n*+*m*, whereby the warranted rate becomes equal to the natural rate of growth. In the special case where *sp*=1 equilibrium between the two is reduced to *r* = *n*+*m*.

Steady state growth with a variable saving ratio and a variable capital-output ratio is shown in Fig.2. *OP* is the production function whose slope measures the marginal productivity of capital (*r*) at any capital-output ratio on a point on *OP*. Equilibrium takes palce where the tangent *WT* touches the *OP* curve at point *A*. The tangent *WT* originates from *W* and not from *O* because savings taking place out of non-wage income *WY*. Point A indicates the rate of profit corresponding to the marginal productivity of capital. In other words, at point *A* labour and capital receive the rewards equal to their marginal productivities. *OW* is the wage rate (the marginal productivity of labour) and *WY* is the profit (the marginal productivity of capital). Thus the steady state equilibrium exists at *A*.

5. Technical Progress

So far we have explained steady state growth without technical progress. Now we introduce technical progress in the model. For this, we take labour augmenting technical progress which increases the effective labour force *L** in the form of a rate of increase in labour productivity.

Assume that the labour force *L* is growing at a constant rate of *n* in year *t*, so that

$$L_t = L_o e^{nt} \qquad \ldots(1)$$

with labour augmenting technical progress, the *effective* labour force L^* is growing at the constant rate of λ in year *t*, so that

$$L^*_t = L^*_o e^{(n+\lambda)t} \quad ...(2)$$

where L^*_O represents the total *effective* labour force in the base period *t*=o embodying all technical progress up to that point in time;

n is the natural growth rate of *effective* labour in the base period;

λ is a constant precentage growth rate of *effective* labour embodied in the base period.

Now the production function for output per worker is

$$q = \frac{Q}{L^*} = \frac{Q}{L_e^{\lambda t}} = f\left(\frac{Q}{L_e^{\lambda t}}\right) = f(k) \quad ...(3)$$

where $\dot{k}$ = *K*/*L**, and the growth rate of $\dot{k}$ (the capital-*effective* labour ratio) is equal to the difference between growth rate of capital stock $\left(\dot{K}\right)$ and the growth rate of *effective* labour $\left(\dot{L}\right)$, *i.e.*

$$\dot{k} = \dot{K} - \dot{L} \quad ...(4)$$

Since $L^* = L_O e^{(n+\lambda)t}$ the growth rate of effective labour L^* is exogenously given as $(n + \lambda)$, so that equation (4) can be written as

$$\dot{k} = \frac{Q}{K} - (n+\lambda)$$

$$= \frac{q}{k} - (n+\lambda)$$

$$\left[\frac{q}{k}=\frac{Q/L}{K/L}=\frac{Q}{K}\right]$$

$= f(k) - (n+\lambda)$ [$\because$ $q = f(k)$ in equation (3) ...(5)

By setting $\dot{k} = O$, we have

$$f(k) = (n+\lambda)k \qquad ...(6)$$

which is the equilibrium condition for steady state growth with technical progress.

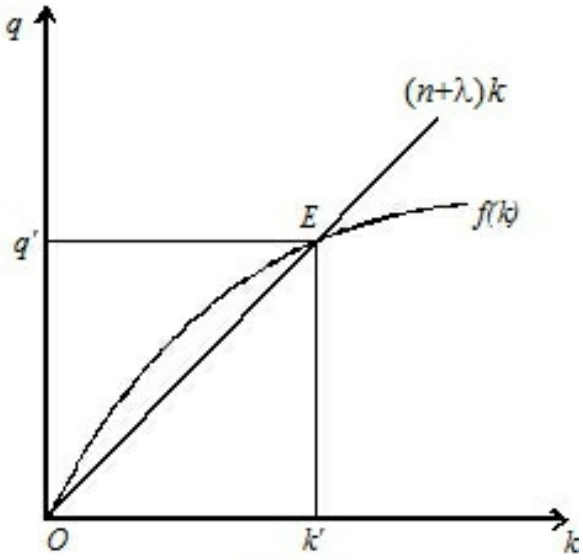


Fig. 3

This is illustrated in Figure 3 where the capital per *effective* worker k is taken horizontally and output per *effective* worker q is taken on the vertical axis. The slope of the ray $(n+\lambda)\ k$ from the origin to point E on the production function $f(k)$ determines the stable equilibrium values k' and q' for k and q repectively at E and the capital used per unit of *effective* labour grows at the rate λ with technical progress.

EXERCISES

1. Explain the basic neo-classical model of growth. Why there is always steady rate of growth in this model?
2. Explain the basic neo-classical growth model.
3. When is an economy in a steady state? Explain the properties of a steady state growth.

CHAPTER 44

THE GOLDEN RULE OF ACCUMULATION

MEANING

Edmund Phelps[1] while explaining the golden age growth lays down the golden rule of accumulation. According to the golden rule of accumulation, per capita consumption is maximised in a golden age when the saving rate equals the profit rate. If the saving rate is less than the profit rate, the per capita consumption is less than the attainable maximum. If, on the other hand, the saving rate is higher than the profit rate, per capita consumption will decline. Thus it is a specific target rate of saving that maximises per capita consumption in a golden age. This is called the *golden rule saving rate* or *optimal saving rate*. If the same amount of consumption is provided to members of each current and future generation, the maximum amount of per capita consumption is called the *golden rule consumption rate*.

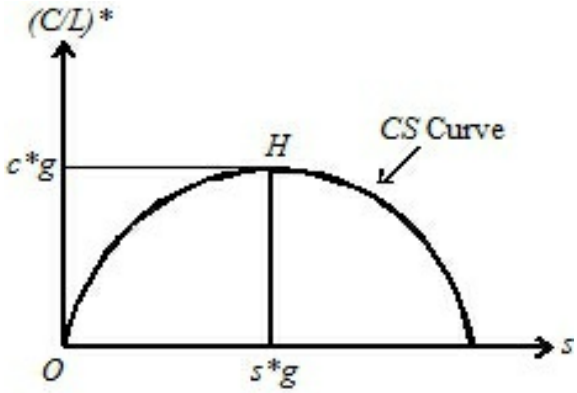


Fig. 1

1. E.S. Phelps, "The Golden Rule of Accumulation : A Fable for Growthman", *A.E.R.*, September 1961.

The golden rule is explained in Figure 1 where golden age per capita consumption, $(C/L)^*$ is taken on the vertical axis and the saving rate s on the horizontal axis. As we move along the CS curve, per capita

consumption increases at low levels of saving and decreases at high levels of saving. It is at the golden rule saving rate s^*g that per capita consumption is maximised at c^*g level of per capita consumption. This is the highest point *H* on the *CS* curve.

GOLDEN AGE AND GOLDEN RULE OF ACCUMULATION

There are several interpretations of Phelps' golden rule of accumulation. We follow Solow's simple explanation.

In a golden age,consumption, output, capital stock and every thing else is growing at the same rate. This growth rate is *g* which is the natural (or equilibrium) growth rate. This rate is independent of the ratio of capital accumulation to output (investment ratio). Corresponding to this investment ratio, there exists at least one capital-output ratio which will be maintained. The existence of a natural growth rate implies that capital and labour are substitutable in such a way that the capital-output ratio can adjust to any value of *s*. Under conditions of equilibrium growth, the capital-output ratio is

$$\frac{K^*}{Q^*} = \frac{s}{g} \quad \text{or } g = s\,(Q^*/K^*)$$

where * (a star) denotes equilibrium value of *K* (capital) and *Q* (output), *s* is saving rate and *g* the golden age growth rate.

The growth rate *g* is not a function of *s*, but it is *s* that determines K^*/Q^*. If we have two golden age paths with equal growth rates, the one with the higher saving rate will have the higher capital-output ratio, and with positive marginal products,the higher level of output. Thus the saving rate affects the *level* at which the economy grows but it does *not* affect its growth rate. This situation is illustrated in Figure 2 where

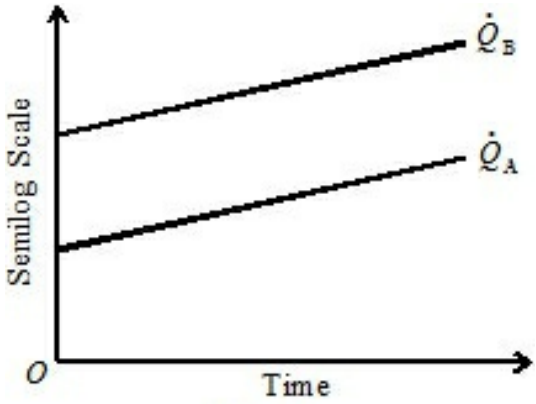


Fig. 2

logarithmatically parallel two golden age paths $\dot{Q}_A$ and $\dot{Q}_B$ pass a given check point, say time zero. They are identical in all respects. But the saving rate on path $\dot{Q}_B$ is higher. This implies that path $\dot{Q}_B$ has a higher capital-output ratio and a higher level of output than path $\dot{Q}_A$.

In deciding which growth path is the best, a generation will look only at the amount of consumption which each path offers. Given a constant *s*, every golden age path is associated with a per capita consumption path $(C/L)^*$ on which consumption grows exponentially at the same rate as output. Under conditions of equilibrium growth, per capita consumption along these parallel paths $(C/L)_1$ *, $(C/L)_2$ * and $(C/L)_3$ * in Figure 3 grows at the indentical rate *g*. Therefore there exists some uniformly highest possible per capita consumption path, such as $(C/L)_3$* than any other consumption path. All generations will naturally prefer this with its corresponding saving rate, to any lower per capita consumption path, such as $(C/L)_2$* or $(C/L)_1$*. Thus the golden rule determines the equilibrium growth path that maximises per capita consumption for *all time*, once the economy reaches that path.

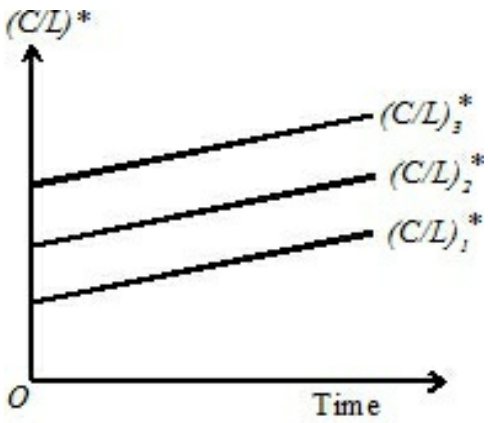


Fig. 3

The higher the saving rate (or investment rate), the higher the level of output. But how much higher? That depends on the marginal product of capital, *r*. Having a little more capital, Δ*K*, at this time would yield *r*Δ*K* more in output. In fact, *r*Δ*K* is the extra output of capital. But all this is not available for consumption. Having a little more capital now commits the economy to some additional investment in the present and in future to keep the slightly lower larger capital stock growing at a fixed rate, *g*. In particular, a little extra capital, Δ*K*, now means that an output of *g*Δ*K* of capital goods is required simply to keep the new little extra capital growing at rate *g*.

If *r*Δ*K* > *g*Δ*K* or *r* > *g*, having a little more capital now will yield some

extra consumption now and for all time. On the other hand, if $r < g$, having a little more capital now would actually siphon off enough productive capacity to the investment sector to reduce consumption.Consumpion is at a maximum when $r = g$ (when the marginal product of capital equals the growth rate). The rate of profit is included in the value of the marginal product of capital. Thus the optimal amount of capital under the golden rule is that which equates the profit rate to the growth rate.[2]

Diagrammatic Representation of the Golden Rule

The golden rule of accumulation is explained in Figure 4. Since the golden rule relates to the maximisation of per capita consumption, we plot output (Q), saving (s), investment (I) and capital stock (K) as ratios to the quantity of labour (L) i.e.

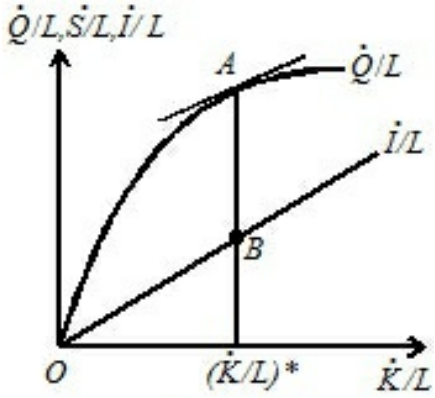


Fig. 3

$$\frac{\dot{I}}{L} = s\left(\frac{\dot{Q}}{L}\right) \qquad [\because I = K] \quad .....(1)$$

or $$\frac{\dot{I}}{L} = g\left(\frac{\dot{K}}{L}\right) \qquad .....(2)$$

where dots over the variables denote golden age magnitudes.

Equation (2) shows that in a golden age per capita investment (or saving) must equal the equilibrium growth rate *times* per capita capital stock.

Therefore, per capita investment is directly proportional to capital-labour ratio. As a result, the $\dot{I} / L$ curve is drawn as a straight line from the origin.

---

2. R.M. Solow, "Comment on the Golden Rule", *R.E.S.*, June 1962.

Now by equating (1) and (2), we get

$$\frac{\dot{I}}{L} = \frac{s}{g} \text{ or } g = s\left(\frac{\dot{Q}}{\dot{K}}\right) \qquad \ldots(3)$$

Equation (3) represents the golden age growth rate which is the slope of the $\dot{I} / L$ curve.

Next the production function.

$$Q = A\,(t)\,(K^{a}, L^{1-a})$$

where $A\,(t)$ is some constant. By holding $t = 0$, and dividing both sides by L, we have

$$\frac{Q}{L} = A\left(\frac{K}{L}\right)^{a}$$

By converting variables in the above equation in relation to the golden age, we have

$$\frac{\dot{Q}}{L} = A\left(\frac{\dot{K}}{L}\right)^{a}$$

which is drawn in the figure as the $\dot{Q} / L$ curve. By partial differentiation, we get the slope of the $\dot{Q} / L$ curve as

$$\frac{d(\dot{Q} / L)}{d(\dot{K} / L)} = a\left(\frac{\dot{Q}}{\dot{K}}\right)$$

Thus a $\left(\dot{Q}/\dot{K}\right)$ is the slope of the $\dot{Q}/L$ cruve which equals the marginal product of capital.

The slope of the $\dot{I}/L$ curve is equal to the golden age growth rate, *g*, and the slope of the $\dot{Q}/L$ curve equals the marginal product of capital, *r* (rate of interest). According to the golden rule of accumulation, the growth rate must equal the rate of interest, $g = r$, in an optimal golden age. This is explained in Figure 4 where $\dot{Q}/L$, $\dot{S}/L$ and $\dot{I}/L$ are taken on the vertical axis and $\dot{K}/L$ on the horizontal axis. The optimal golden age, as per the golden rule of accumulation, is at the equilibrium capital-labour ratio $\left(\dot{K}/L\right)^{*}$ where the vertical distance between the curves $\dot{Q}/L$ and $\dot{I}/L$ is the greatest. It is *AB* in the figure. Thus the golden rule of accumulation is satisfied when the vertical difference between the two curves $\dot{Q}/L$ and $\dot{I}/L$ is at the maximum. It is here that the per capita consumtion is maximised.

EXERCISES

1. What do you mean by the Golden Rule of Accumulation? How does an economy reach the golden rule of accumulation?
2. What is Golden Age Path? “On the optimum natural growth path the investment ratio and the profit ratio are constant and equal.” (E. Phelps) Explain this statement.

PART-VII

MACROECONOMIC POLICIES

CHAPTER 45

MACROECONOMIC POLICIES

MEANING

Macroeconomic Policy refers to the instruments by which a government tries to regulate or modify the economic affairs of the country in keeping with certain objectives. It "attempts to assess the behaviour of the economy as a whole and to seek ways in which its aggregate performance might be improved."[1] To Shapiro, "Macroeconomic policy means that government does such things as slow down or speed up the growth rate of the money supply, raise or cut tax rates and/or the level of government spending, and impose or withdraw controls on prices and wages."[2] Macroeconomic policy is achieved through certain instruments and objectives. Its two main instruments are monetary and fiscal policy and its four major objectives are full employment, price stability, economic growth, and balance of payments equilibrium.[3] In the present chapter we study the objectives of macroeconomic policy and the problems that arise when these objectives come into conflict with each other.

1. G.K. Shaw, *An Introduction to the Theory of Macroeconomic Policy*, p.1.

2. Edward Shapiro, *op. cit*., p. 546.

POLICY TARGETS AND INSTRUMENTS

The policy targets are the specific values which a government attaches to its various objectives of macroeconomic policies. For instance, the government may have the following policy objectives : (1) to achieve full

employment at the rate of 3 per cent unemployment; (2) to achieve price stability at an annual inflation rate of 5 per cent per annum; and (3) to attain the growth rate of 5 per cent per annum for the economy. Thus the policy targets of the government are 3 per cent unemployment rate, 5 per cent inflation rate and 5 per cent growth rate per year. On the other hand, policy instruments are those exogenous variables that can be directly influenced by the government. The government can influence macroeconomic policies by such instruments of monetary policy as bank rate, changes in reserve ratios, open market operations, selective credit controls, etc. Similarly, it can use such fiscal policy instruments as tax rates, budgetary policy, compensatory fiscal policy, etc.

OBJECTIVES OF MACROECONOMIC POLICY

The following are the objectives of macroeconomic policy.*

Full Employment

Full employment has been ranked among the foremost objectives of economic policy. But there is no unanimity of views on the meaning of full employment. Prof. Ackley regards it as a "slippery concept." But the credit for popularising it goes to Keynes and since the Second World War it has been accepted as one of the important goals of macroeconomic policy.

The classical economists always believed in the existence of full employment in the economy. To them full employment was a normal situation and any deviation from this was regarded as something abnormal. According to Pigou, the tendency of the economic system was to automatically provide full employment in the labour market. Unemployment resulted from the rigidity in the wage structure and interference in the working of the market system in the form of trade union legislation, minimum wage legislation, etc. Full employment existed "when everybody who at the running rate of wages wishes to be employed." Those who are not prepared to work at the existing wage rate are not unemployed in the Pigovian sense because they are voluntarily unemployed. There is, however, no possibility of involuntary

unemployment in the sense that people are prepared to work but they do not find work. According to Pigou, “With perfectly free competition—there will always be at work a strong tendency for wage rates to be so related to demand that everybody is employed.” However, this classical view on full employment is consistent with some amount of frictional, voluntary, seasonal or structural unemployment.

---

3. There appears to be near unanimity over these objectives, although certain economists also include balanced regional development and just distribution of income among the goals of economic policy. Interested readers may refer to K.E. Boulding, *Principles of Economic Policy*, 1958; A.K. Dasgupta and A.J. Hagger, *The Objectives of Macroeconomic Policy*, 1971.

* The objectives of macroeconomic policy are also the objectives of monetary policy and fiscal policy.

According to Keynes, full employment means the absence of involuntary unemployment. In other words, full employment is a situation in which everybody who wants to work, gets work. Full employment so defined is consistent with frictional and voluntary unemployment. Keynes assumed that “with a given organisation, equipment and technique, real wages and the volume of output (and hence of employment) are uniquely co-related so that, in general, an increase in employment can only occur to the accompaniment of a decline in the rate of wages.” To achieve full employment Keynes advocated increase in effective demand to bring about reduction in real wages. Thus the problem of full employment is one of maintaining adequate effective demand. Keynes gave an alternative definition of full employment at another place in his *General Theory* thus : “It is a situation in which aggregate employment is inelastic in response to an increase in the effective demand for its output.” It means that the test of full employment is when any further increase in effective demand is not accompanied by any increase in output. Since the supply of output becomes inelastic at the full employment level, any further increase in effective demand will lead to inflation in the economy. Thus the Keynesian concept of full employment involves three conditions: (i) reduction in the real wage rate; (ii) increase in effective demand; and (iii) inelastic supply of outut at the level of full employment.

According to Professor W.W. Hart attempting to define full employment raises many people’s blood pressure. Rightly so, because there is hardly

any economist who does not define it in his own way. Lord Beveridge in his book *Full Employment in a Free Society* defined it as a situation where there were more vacant jobs than unemployed men so that normal lag between losing one job and finding another will be very short. By full employment he does not mean zero unemployment which means that full employment is not always full. There is always a certain amount of frictional unemployment in the economy even when there is full employment. He estimated frictional unemployment of 3% in a full employment situation for England. But his pleading for more vacant jobs than the unemployed cannot be accepted as the full employment level.

According to the American Economic Association Committee, "Full employment is a situation where all qualified persons who want jobs at current wage rate find full-time jobs." It does not mean unemployment is zero. Here again, like Beveridge, the Committee considered full employment to be consistent with some amount of unemployment.

Individual economists may, however, continue to differ over the dkefinition of full employment, but the majority has veered round the view expressed by the U.N. experts on *National and International Measures for Full Employment* that "full employment may be considered as a situation in which employment cannot be increased by an increase in effective demand and unemployment does not exceed the minimum allowances that must be made for the effects of frictional and seasonal factors." This definition is in keeping with the Keynesian and Beveridgian views on full employment. It is now agreed that full employment stands for 96 to 97 per cent employment, with 3 to 4 per cent unemployment existing in the economy due to frictional factors. For achieving full employment, a judicious mix of monetary-fiscal policies is used.

Price Stability

One of the policy objectives of monetary and fiscal policy is to stabilise the price level. Both economists and laymen favour this policy because fluctuations in prices bring uncertainty and instability to the economy. Rising and falling prices are both bad because they bring unnecessary loss to some and undue advantage to others. Again they are associated with business cycles. So a policy of price stability keeps the value of money

stable, eliminates cyclical fluctuations, brings economic stability, helps in reducting inequalities of income and walth, secures social justice and promotes economic welfare.

However, there are certain difficulties in pursuing a policy of stable price level. The first problem relates to the type of price level to be stabilised. Should the relative or general price level be stabilised, the wholesale or retail, of consumer goods or producer goods? There is no specific creiteria with regard to the choice of a price level. Halm suggests, "The compromise solution would be to try to stabilise a price level which would include consumers' goods prices as well as wages." But this will necessitate increase in the quantity of money but not by as much as is implied in the stabilisation of consumers' goods prices. Second, innovations may reduce the cost of production but a policy of stable prices may bring larger profits to producers at the cost of consumers and wage earners. Again, in an open economy which imports raw materials and other intermediate products at high prices the cost of production of domestic goods will rise. But a policy of stable prices will reduce profits and retard further investment. Under the circumstances, a policy of stable prices is not only inequitable but also conflicts with economic progress.

Despite these drawbacks, the majority of economists favour a policy of stable prices. But the problem is one of defining price stability. Price stability does not mean that prices remain unchanged indefinitely. "Comparative prices will change as fluctuating tests alter the composition of demand, as new products are developed and as cost reducing technologies are introduced. Differential price changes are essential for allocating resources in the market economy. Since modern economies tend to exhibit fairly rigid downward inflexibility of prices, differential price changes can only be attained by gradual increases in the aggregate price level over the long run."[4] Further, prices may have to be changed if costs of imported goods increase or if taxation policy leads to the rise in the domestic cost of production. Therefore, economists agree that a moderate rate of inflation is desirable for the healthy growth of the economy which may not be in excess of 2% to 3%. Such price stability can be maintained by following counter-cyclical monetary and fiscal policies.

Economic Growth

One of the most important objectives of macroeconomic policy in recent years has been the rapid economic growth of an economy. Economic growth is defined as "*the process whereby the real per capita income of a country increases over a long period of time*."[5] Economic growth is measured by the increase in the amount of goods and services produced in a country. A growing economy produces more goods and services in each successive time period. Thus growth occurs when an economy's productive capacity increases which, in turn, is used to produce more goods and services. In its wider aspect, economic growth implies raising the standard of living of the people, and reducing inequalities of income distribution. All agree that economic growth is a desirable goal for a country. But there is no agreement over "the magic number," viz., the annual growth rate which an economy should attain.

---

4. A.K. Dajgupta and A.J. Hagger, *Ibid*.

5. G.M. Meier, *Leading Issues in Development Economics, p. 20.*

Generally, economists believe in the possiblity of continual growth. This belief is based on the presumption that innovations tend to increase productive technologies of both capital and labour over time. But there is every likelihood that an economy might not grow despite technological innovations. Production might not increase further due to the lack of demand which may retard the growth of the productive capacity of the economy. The economy may not grow further if there is no improvement in the quality of labour in keeping with the new technologies.

However, policy makers do not take into consideration the costs of growth. Growth is not limitless because resources are scarce in every economy. All factors have opportunity cost. To produce more of one particular product will mean reduction in that of the other. New technologies lead to the replacement of old machines which become useless. Workers are also displaced because they cannot be fitted in the new technological set-up immediately. Moreover, rapid growth leads to urbanisation and industrialisation with their adverse effects on the pattern of living and environment. People have to live in squalor and slums. The

environment becomes polluted. Social tensions develop. “But growth has other more basic effects on our environment, and, today, people are not so sure that unrestricted growth is worth all its costs, since the price in terms of change in, deterioration of, or even destruction of the environment is not yet fully known. What does seem clear, however, is that growth is not yet fully known. What does seem clear, however, is that growth is not going to be halted because of environmental problems and that mankind must learn to cope with the problems or face the consequences.”

Monetary and fiscal policies contribute towards growth by helping to maintain stability of prices. By moderating economic fluctuations and avoiding recessions, these policies help in achieving the growth objective. Since rapid and variable rates of inflation discourage investment and adversely affect growth, these policies help in controlling hyper-inflation. And growth can be promoted by a judicious mix of monetary-fiscal policies. So monetary and fiscal policies should be such as to encourage investment, and control economic fluctuations in order to promote growth.

Balance of Payments

Another objective of macroeconomic policy since the 1950s has been to maintain equilibrium in the balance of payments. The achievement of this goal has been necessitated by the phenomenal growth in the world trade as against the growth of international liquidity. It is also recognised that deficit in the balance of payments will retard the attainment of other objectives. This is because a deficit in the balance of payments leads to a sizeable outflow of gold. But “it is not clear what constitutes a satisfactory balance of payments position. Clearly a country with a net debt must be at a surplus to repay the debt over a reasonably short period of time. Once any debt has been repaid and an adequate reserve attained, a zero balance maintained over time would meet the policy objective. But how is this satisfactory balance to be achieved on the trading account or on the capital accout ? ...The capital account must be looked upon as fulfilling merely a short term emergency role in times of crisis”[6]

Again, another problem relates to the question: What is the balance of payment target of a country ? It is where imports equal exports. But, in a

practice, a country whose current reserves of foreign exchange are in adequate will have a mild export surplus as its balance of payments target. But when its reserves become satisfactory, it will aim at the equality of imports and exports. This is because an export surplus means that the country is accumulating foreign exchange and it is producing more than it is consuming. This will lead to low standard of living of the people. But this cannot last long because some other country must be having import surplus and in order to avoid it, it would impose trade restrictions on the export surplus country. So the attainment of a balance of payments equilibrium becomes an imperative goal of economic policy in a country.

A deficit in the balance of payments of a country can be wiped out with restrictive monetary and fiscal policies, by reducing imports and encouraging exports, and by devaluation of the currency.

CONFLICTS OR TRADE-OFF IN POLICY OBJECTIVES

The four policy objectives discussed above are not complementary to each other. Rather, they conflict with one another. If a government tries to fulfil one objective, some other objective moves away. It has to sacrifice one objective in order to attain the other. It is, therefore, not possible to fulfil all these policy objectives simultaneously. We discuss below conflicts between different policy objectives.

Full Employment and Economic Growth : Okun's Law

The majority of economists hold the view that there is no inherent conflict between full employment and economic growth. Full employment is consistent with 4 per cent unemployment in the economy. So the relationship between full employment and economic growth boils down to a trade-off between unemployment and growth. Periods of high growth are associated with low level of unemployment, and periods of low growth with rising unemployment.

In 1961, Aurther Okun[7] established a relationship between real GNP and changes in the unemployment rate. This relationship has come to be known as Okun's Law. This law states that *for every three percentage*

points growth in real GNP, unemployment rate declines by one percentage point every year. This is illustrated in Figure 1 where the curve U represents unemployment and curve G the real growth of an economy for a few years. To begin with, suppose that the economy is growing at 3 per cent with an unemployment rate of 4 per cent. During the year 1970, when the real GNP increases by 4.5 percent (from 3 per cent to 7.5 per cent), the unemployment rate falls by 1.5 per cent (from 4 per cent to 2.5 per cent). In the next year 1971, the growth rate of the economy falls to zero and the unemployment rate rises to 5 per cent. In the subsequent year 1972, the real growth rate increases to 3 per cent and the unemployment rate declines to 4 per cent.

---

6. G.K. Shaw, *op.cit*., p.9.

7. A.M. Okun, "Potential GNP : Its Measurement and Significance," in W.L. Smith and R.L. Teigen (eds.), *op.cit.*

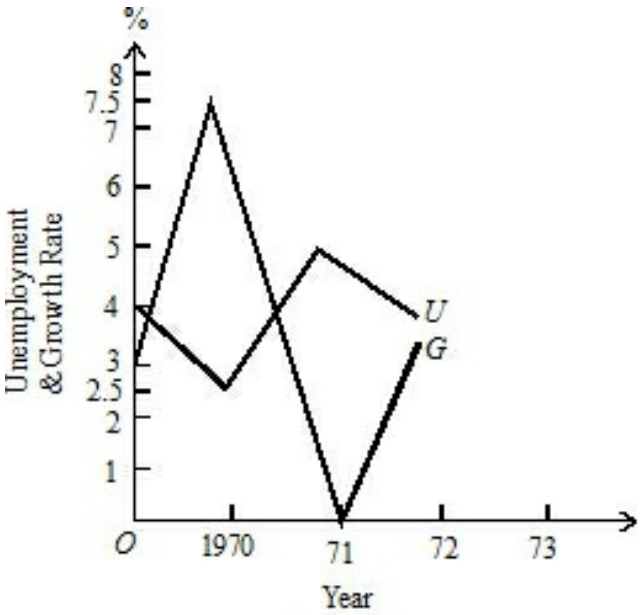


Fig. 1

However, certain economists argue that the unemployment rate increases as the growth rate rises. Economic growth leads to reallocation of resources in the economy whereby there is change in the type and quantity of labour demanded. There is shift in the demand for labour from one sector of the economy to the other. As workers are trained for specific jobs, they are displaced when the demand for the products of particular industries falls. This creates unemployment. This is particularly so when growth is the result of technological innovations which are labour-saving and require more qualified and skilled workers. Thus unskilled workers are the worst sufferers because they are thrown out of jobs with automation. Employment can, however, incease with growth if demand expands faster than the productivity of labour. If demand is increasing at 3 per cent per annum and the productivity is

increasing at 4 per cent per year, the output will expand but employment will decline. Under the circumstances, the government should adopt such monetary and fiscal policy which should increase the overall demand in the economy.

Economic Growth and Price Stability

There is conflict between the goals of economic growth and price stability. The rise in prices is inherent in the growth process. The demand for goods and services rises as a result of stepping up of investments on a large scale and consequent increase in incomes. This leads to inflationary rise in prices, especially when the level of full employment is reached. In the long run, when new resources are developed and growth leads to the production of more commodities, the inflationary rise in prices will be checked. But the rise in prices will be there with the growth of the economy and it will be moderate and gradual.

Full Employment and Price Stability

One of the objectives of macroeconomic policy in the 1950s was to have full employment with price stability. But the studies of Philips, Samuelson, Solow and others in the 1960s established a conflict between the two objectives. These findings are explained in terms of the Philips curve. They suggest that full employment can be attained by having more inflation and that price stability can be achieved by having unemployment to the extent of 5 to 6 per cent. Economists do not find any conflict between unemployment and price stability. They hold that so long as there are unemployed resources, there will be price stability. Prices start rising only when there is full employment of resources. This is illustrated in Figure 2 where the percentage of resources unutilized (or unemployed) are taken on the horizontal axis and the percentage change in price level is

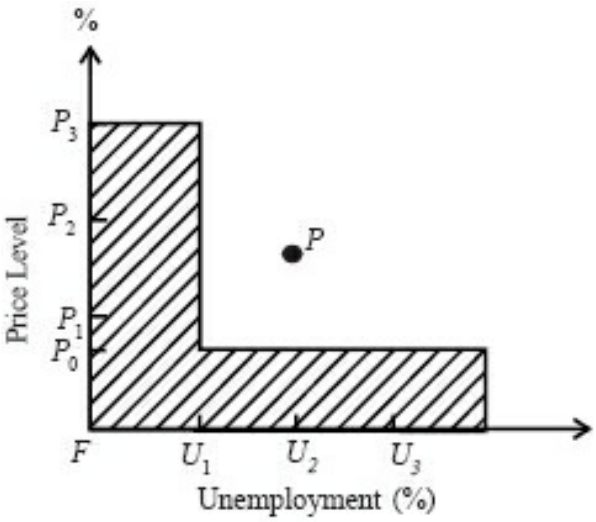


Fig. 2.

taken on the vertical axis. Thus each point indicates the percentage of resources unemployed along with the price level. According to this theory, so long as resources U_1, U_2 and U_3 are unemployed, the price level remains constant at P_O. It is only when the economy reaches the full employment level F, prices rise from P_O to P_1 to P_2 to P_3 with successive increases in demand. Thus there is no conflict between unemployment and stable prices, as shown by the shaded area of the figure. However, the macro policy implications of such a relationship are that there can be no conflict between full employment and price stability so long as the economy is in the shaded area. This is because when there is full employment, resources are not in excess supply and if the government controls the excess demand through appropriate monetary and fiscal policy, there will be stability of the price level. But if the economy happens to be at point P, which may be taken to be a point on the Phillips curve, there will be conflict between the objectives of full employment and price stability.

Full Employment and Balance of Payments

There is a major policy conflict between full employment and balance of payments. Full employment is always related to balance of payments deficit. In fact, the problem is one of maintaining either internal balance or external balance. If there is a balance of payments deficit, then a policy of reducing expenditure will reduce imports but it will lead to increase in unemployment in the country. If the government raises aggregate expenditure in order to increase employment, it will increase the demand for imports thereby creating disequilibrium in the balance of payments. It is only when the government adopts expenditure-switching policies such as devaluation that this conflict can be avoided but that too temporarily.

Price Stability and Balance of Payments

There appears to be no conflict between the objectives of price stability and balance of payments in a country. Fiscal and monetary policies aim at controlling inflation to discourage imports and encourage exports and thus they help in attaining balance of payments equilibrium *imports and encourage exports and thus they help in attaining balnce payments*

equilibrium. However, if the government tries to remove unemployment and allows some inflation within the economy, there appears a conflict between these two objectives. For a rise in the price level will discourage exports and encourage imports, thereby leading to disequilibrium in the balance of payments. But this may not happen if prices also rise by the same rate in other countries of the world.

PROBLEM OF COORDINATION OF MACROECONOMIC POLICY OBJECTIVES*

We have seen above that there are four policy goals which often conflict with each other. The problem is one of achieving them simultaneously. Full employment, economic growth and price stability are the major objectives of economic policy. They are essential for the internal balance of the economy. But balance of payments equilibrium is also an essential policy objective, because a disturbance in the balance of payments has serious effects on growth, employment and prices. This objective, therefore, requires external balance.

The theory of economic policy has centred around two distinct problems. First, the relation between the number of policy objectives and the number of policy instruments; and second, the assignment of policy instruments to the realisation of the objectives.

The first problem has been explained by Trevor Swan and the second by R.A. Mundell which are discussed below.

THE SWAN MODEL

Trevor Swan[8] takes two targets relating to full employment and balance of payments and two policy instruments expenditure-reducing and expenditure-switching policies in order to achieve coordination between targets and instruments.

To achieve the two targets, a judicious combination of expenditure-reducing (internal policies) and expenditure-switching (external policies)

instrument is needed. For instance, if the economy is already at the full employment level, a policy of devaluation may cause inflation within the economy. So expenditure-switching policy of devaluation must be accompanied by expenditure-reducing policies of tighter fiscal and monetary controls to maintain full employment and balance of payments equilibrium.

Assumptions

The Swan model is based on the assumptions that (1) there are no trade restrictions; and (2) there are no capital movements (3) productivity, terms of trade and other financial transfers are given.

---

* This portion may be read after studying monetary and fiscal policies.

8.. "Long-Run Problems of Balance of Payments," in R.E. Caves and H.G. Johnson, *op.cit.*, Ch. 27.

Explanation

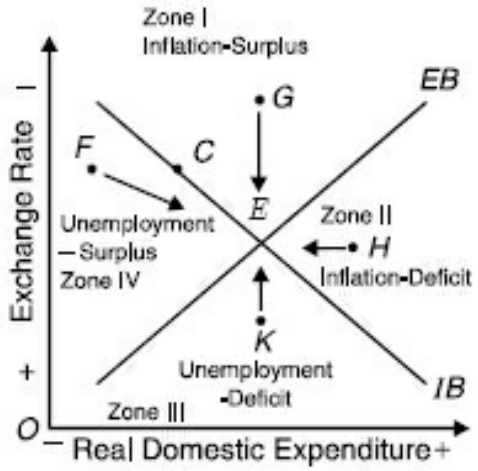


Fig. 3

Given these assumptions, the model is explained in Figure 3 where the horizontal axis measures real domestic expenditure, and the vertical axis the exchange rate. A movement to the left (towards *O*) on the horizontal axis means the use of expenditure-reducing policy, and a movement upwards along the vertical axis means the use of expenditure-switching policy. *IB* is the internal balance curve which represents a situation of full employment and stable prices. It represents the various combinations of exchange rates and real domestic expenditure. The *IB* curve is negatively sloped which shows that an increase in domestic absorption (real expenditure) must be balanced by a decrease in the exchange rate of the country to reduce its trade balance in order to maintain full employment with price stability. Obviously, *points to the right and above the IB curve relate to inflation* or *over full employment,* and *points to the left and below the curve refer to recession or unemployment.*

The *EB* curve represents the external balance where exports equal imports, in the absence of capital movements. So external balance occurs when net exports equal zero. This curve slopes upwards from left to right which shows that for the economy to remain in external balance, devaluation must be balanced by an increase in domestic expenditure (Devaluation will improve the country's trade balance by encouraging exports and discouraging imports, and the increase in real domestic expenditure will increase the country's imports sufficiently). Obviously, *points above the EB curve refer to a surplus and points below the curve relate to a deficit in the balance of payments.*

The point where the *IB* curve intersects the *EB* curve represents the point of "bliss" where the economy is simultaneously in internal and external balance. *E* is such a point in Figure 3 where the exchange rate and real domestic expenditure are in equilibrium. If the economy is not at point *E* , it is in disequilibrium. According to Swan, "The two curves of internal balance and external balance divide existence into four zones of economic unhappiness." The four zones of disequilibrium are :

Zone I : Inflation and *BOP* surplus.

Zone II : Inflation and *BOP* deficit.

Zone III : Unemployment and *BOP* deficit.

Zone IV : Unemployment and *BOP* surplus.

To explain the type of policy measures which may be required to achieve internal and external balance simultaneously, we take disequilibrium positions in the four zones.

Take point *G* in Zone I where a *BOP* surplus is combined with inflation. In this situation, the exchange rate should be appreciated to correct the *BOP* surplus and expenditure be reduced to combat inflation. But reduction in expenditure would increase the *BOP* surplus. This represents the "dilemma zone" because no uniform policy can be adopted.

Similarly, point *K* in Zone III where unemployment and *BOP* deficit exist

require increase in domestic expenditure to reduce unemployment and depreciation of exchange rate to correct *BOP* deficit. But an expansionary fiscal policy increases income and demand, and thus widens *BOP* deficit. This is again the dilemma zone.

Take point *H* in Zone II where inflation is combined with *BOP* deficit. Inflation should be combated with reduction in domestic expenditure which would also reduce *BOP* deficit along with depreciation of exchange rate. Ultimately, the economy will move towards the equilibrium position *E*.

Finally, move to point *F* in Zone IV where *BOP* surplus is combined with unemployment. Here appreciation of exchange rate and increase in domestic expenditure will move the economy towards internal and external balance at *E*. Zones II and IV are "simple zones".

Conclusion. The above discussion reveals that if the economy is on neither the *IB* curve nor the *EB* curve, it is in one of the four zones. When the economy follows expenditure changing monetary and fiscal policies simultaneously in dilemma Zones I and III to achieve internal balance (one target), it moves away from external balance (the other target). To solve this dilemma, Tinbergen developed the rule known as the *Tinbergen Principle* which leads to the assignment problem. The solution to this problem has been suggested by Mundell which we discuss below in detail.

THE ASSIGNMENT PROBLEM : THE MUNDELLIAN MODEL

The Assignment Problem

Jan Tinbergen[9] was the first economist to lay down that the number of policy instruments must be equal to the number of objectives. If there are more objectives than policy instruments it means that there are not enough tools to achieve the policy objectives. The system is undetermined. On the other hand, if the number of policy instruments is more than the number of objectives, then there is not one combination of tools and objectives that will solve the problem, but any number. The system is over-determined. Thus the number of policy tools must equal the number of

targets for economic policy to be successful. This has come to be known as the *Tinbergen Principle* or the fixed targets approach.

In order to achieve given objectives with the same number of policy instruments, the second problem of the assignment of instruments to targets arises. The formulation of the assignments problem will eventually lead to equilibrium values of the objectives, despite lack of co-ordination between them. Thus the assignment probelm relates to the assignment of instruments to targets. The solution to the assignment problem has been suggested by Robert Mundell by the *Principle of Effective Market Classification.*[10]

The Mundellian Model

Mundell discusses the case of relationship between two instruments and two targets. The two instruments are monetary policy represented by interest rate and fiscal policy represented by government expenditure. The two objectives or targets are full employment (internal balance) and balance of payments equilibrium (external balance). The *assignment rule is to assign monetary policy to the objective of external balance and fiscal policy to internal balance.*

---

9. J.Tinbergen, *On the Theory of Economic Policy,* 1952.

10. R.A. Mundell, "The Appropriate Use of fiscal and Monetary Policy for Internal and External Stability," *I.M.F. Staff Papers, 1962. International Economics,* 1968.

Its Assumptions

The Mundellian model is based on the following assumptions:

1. Monetary policy is related to changes in interest rate.

2. Fiscal policy is related to deficit or surplus budget.

3. Exports are exogeneously given.

4. Imports are a positive function of income.

5. International capital movements respond to domestic interest rate changes.

The Model

Given these assumptions, Mundell states that "in countries where employment and balance of payments policies are restricted to monetary and fiscal instruments, monetary policy should be reserved for attaining the desired level of the balance of payments, and fiscal policy for preserving internal stability under the conditions assumed here." If monetary and fiscal policies are adjusted smoothly and continuously without long time lags, the assignment rule can work very well. In some cases, it leads straight to the target, while in others it may worsen the other problems temporarily. But ultimately it will achieve the target. This is Mundell's principle of effective use of monetary and fiscal policy for internal and external stability, according to which an instrument should be matched with the target on which it exerts the greatest relative influence. He calls it the *Principle of Effective Market Classification.*

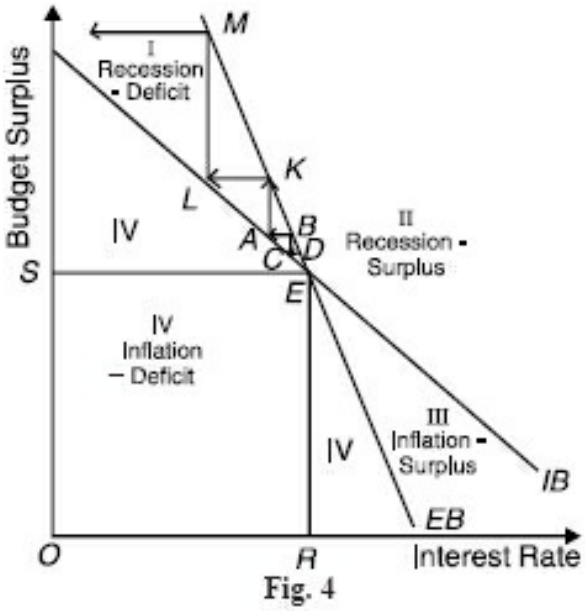


Fig. 4

In Mundell's slightly modified Figure 4, the horizontal axis measures interest rate (monetary policy) and the vertical axis budget surplus (fiscal policy). *IB* is the internal balance line and *EB* the external balance line. The *IB* line represents full employment. It is negatively sloped because a reduction in budget surplus* must be balanced by an increase in interest rate** in order to maintain full employment. There is inflation below this line *IB* (Zone III and IV), and recession above it (Zone I and II). On the other hand, line *EB* gives all points of equilibrium in the balance of payments. It is also negatively sloped because a reduction in the budget surplus must be counteracted by increase in interest rate. There is deficit in the balance of payments below

this line (Zone I and IV), and surplus above this line (Zone II and III). The *EB* line is steeper than the *IB* line because an increase in interest rate in order to balance an expansionary fiscal policy (increase in budget deficit or reduction in budget surplus) induces a short-term capital inflow for an external balance. The more responsive capital movements are to interest rate changes, the steeper is the *EB* line relative to the *IB* line. This makes monetary policy relatively more effective for maintaining external balance.

---

* Reduction in budget surplus means increase in government expenditure. It is contractionary fiscal policy.

** Increase in interest rate implies reduction in the money supply. It is contractionary monetary policy.

Figure 4 illustrates external and internal balance and the role played by monetary and fiscal policy in maintaining equilibrium between the two at point *E* with *OR* interest rate and *OS* budget surplus. The two policy measures will take the economy to the equilibrium point *E* in inflation-deficit and recession-surplus situations. But there are policy conflicts in inflation-surplus (Zone III) and recession-deficit (Zone I). Suppose the economy is at point *A* in Zone IV where there is full employment within the economy and deficit in the balance of payments. To remove balance of payments deficit, the *monetary authority* acts first by increasing the interest rate by *AB* in order to reduce the money supply. The reduction in money supply will reduce demand for goods and this will, in turn, decrease imports, and restore equilibrium in the balance of payments at *B*. But here the economy is having recession and unemployment. To correct these and to have internal balance, budget surplus will have to be reduced by *BC*. But at *C*, there is again deficit in the balance of payments which necessitates further increase in interest rate by *CD* for reducing the money supply. At *D* the internal balance is again disturbed leading to a further reduction in budget surplus. This process of reduction in money supply followed by reduction in budget surplus will ultimately lead the economy to the equilibrium point *E* where there is simultaneous internal and external balance. Thus the use of monetary policy for external balance and fiscal policy for internal balance will lead to equilibrium in Zones II and IV.

On the other hand, if *budget surplus is used to remedy the deficit in balance of payments and monetary policy to correct recession and unemployment, there would be neither external balance nor internal balance*. Starting from point *A*, an increase in budget surplus would move the economy to *K* where the external balance is achieved but there is recession and unemployment in the economy. To remedy it, the interest rate is reduced by *KL* for increasing the money supply. But at *L* deficit in the balance of payments rises over its previous level. This will require a still greater budget surplus by *LM*. This will necessitate still larger reduction in interest rate to remove recession and unemployment. In this way, the economy would move further and further away from point *E* and there would not be simultaneous internal and external balance in Zone I. In this case, the assignment rule leads to explosive instability because the two policies are badly coordinated. Similar will be the case in Zone III.

Thus monetary policy should be assigned to the objective of external balance and fiscal policy to that of internal balance. However, the assignment rule can work only if monetary and fiscal policy can be adjusted smoothly and continuously without long lags before their effects are visible. This is Mundell's principle of effective use of monetary and fiscal policy for internal and external stability, according to which an instrument should be matched with the target on which it exerts the greatest relative influence. He calls it the *Principle of Effective Market Classification.*

Criticisms of Mundell's Model

There are several shortcomings of Mundell's analysis.

1. Unrealistic Assumptions. This model assumes that the authorities know about the *IB* and *EB* curves, the zone in which the economy is operating, and the extent to which the economy is away from both internal and external balance so that appropriate monetary and fiscal policy can be applied. It also presupposes that they know the quantitative results which are expected from the application of each policy. But it is not possible to accurately estimate the degree of disequilibrium due to lack of data about them. Accordingly, the policy changes may not be appropriate to the type of disequilibrium.

2. Ignores Stagflation. This analysis overlooks the situation of unemployment and inflation. This is unrealistic because this phenomenon, known as stagflation, is found in almost all developed countries.

3. Neglects Other Factors. This analysis considers only differences in interest rates as the cause of capital movements and neglects other factors such as exchange rate variations. Moreover, it is not possible that a persistent deficit may be financed by means of capital movements.

4. Practical Constraints. Monetary and fiscal policies operate under certain practical constraints. Due to political reasons, some governments are unable to follow a restrictive fiscal policy and a monetary policy of high interest rates. Even if such policies can be started, they may not be successful because capital flows may not be interest-sensitive.

5. Cannot Correct Current Account Deficit. The prescribed policy mix may be unable to correct a current account deficit. Since the policy mix affects both the capital flows and imports, it can only ensure that a negative trade balance is offset by a positive capital flow, and vice versa.

6. Not a True Adjustment Mechanism. The monetary-fiscal mix is not a true adjustment mechanism. It is just a palliative. It does not adjust the balance of payments but simply stabilises it. Capital flows only fill the gap between autonomous demand and supply of foreign exchange, leaving prices and incomes unchanged.

7. Debt-Servicing Requirements not Considered. This analysis does not take into account the debt-servicing requirements that a continuous capital inflow would have on the current account of the balance of payments when the domestic interest rate is raised.

8. Retards Capital Formation. When the interest rate is raised through monetary policy, it will lead to a decrease in investment at home. This must be accompanied either by an increase in government expenditure or by tax reductions or by a combination of both. Such a monetary-fiscal mix wastes the economy’s savings by diverting them into debt-financed government expenditure which retards capital formation. According to Johnson, this raises the problem of “inefficiency versus efficiency in the

use of domestic saving potential."

9. Conflicting Policies. There is the possibility of conflicts between the prescribed policy mixes among governments of different countries. According to Johnson, it is a difficult and highly complicated process to arrive at the right combination of monetary and fiscal policies in all countries simultaneously. If, however, such a combination is worked out by trial and error, it may lead away from rather than toward equilibrium.

10. Long Time Lags. The model assumes that there are no long time lags for the operation of monetary-fiscal policies. But both monetary and fiscal lags are quite long and they retard the process of simultaneous equilibrium of internal and external balance.

RULES VS DISCRETION IN ECONOMIC POLICY

In macroeconomic policy, there is a controversy over the issue of "rules vs discretion". The issue is whether monetary and fiscal authorities should conduct policies in accordance with predetermined rules that lay down how their policy variables will be determined in all future economic situations, or whether they should be allowed to use their discretion in determining the policy variables as situations arise.

As the economy is inherently unstable due to frequent aggregate demand and supply shocks, these shocks lead to fluctuations in output, employment and prices. Keynesians led by Tobin favour discretionary or activist policy while monetarists led by Friedman oppose activist policy and favour passive policy based on rules.

According to Keynesians, the economy is subject to many exogenous shocks from factors like changes in expectations, political events, international events such as oil crisis, war, etc. But usually they are money shocks, demand shocks or price shocks. These exogenous shocks lead to irregular cyclical fluctuations of varying intensity. Fluctuations being unpredictable, they bring uncertainty in the economy thereby adversely affecting investment and potential output. So they lead to recessions resulting in high unemployment, fall in incomes and hardships to people.

Morever, Keynesian analysis is based on the premise that the private sector is inherently unstable and is subject to frequent disturbances in aggregate output, income and employment. If left to itself, it will cause deviation from real income, output and employment from the full employment level. Therefore, Keynesians argue that government should follow activist fiscal and monetary policies to stabilise the economy.

To Keynesians, stabilisation policy means “leaning against the prevailing economic wind.” It should stimulate the economy in recession and slow it when overheated in boom. This requires deliberate changes in tax rates and government expenditures, and tight or easy monetary policy during booms or recessions. Thus activist or discretionary policy is flexible in which government, monetary or other authorities decide what appropriate policy should be adopted in current circumstances. It is formulated on a case-by-case on yearly basis in which there is no commitment about future policy.

Keynesian discretionary policy has been criticised by monetarists who regard the private sector as inherently stable and consequently the economy is basically stable. When disturbed by some change in basic conditions, it will quickly revert to its long run path via relative price changes. So discretionary stabilisation policy is not required. If such a policy is implemented, it will increase rather than diminish fluctuations in output, income and employment. They do not favour “fine tuning” the economy and are critical of counter-cyclical discretionary policy. Such a policy involves long, variable and uncertain time lags that make stabilisation policy ineffective and destabilising. So the best policy is a passive one which follows a “fixed long-run rule”. Under such a rule, the money supply should grow at a constant rate of 3 per cent per year and government’s fiscal deficit is fixed at 3 per cent of GDP per year. But economists do not agree with monetarists on fixed long-run rules because rules do not grow at the same rate. Economic environment may change and past rules may no longer hold. Therefore, government should have the discretion to change the rules. There should be flexibility of instruments for maintaining stability of targets. For instance, money supply and government spending are instruments of stabilisation policy that affect such targets as prices, output, income and employment. Moreover, given

that the economy and the knowledge about it are both changing overtime, there is no need for laying down permanent policy rules that would tie down the hands of monetary and fiscal authorities. Thus the practical issue in the controversy over rules versus discretion becomes whether the policy makers should announce in advance what policies they will follow in the foreseable future.

Expectations, the Lucas Critique and New Classical Stabilisation Policy

New classicals do not agree with Keynesians and monetarists about macro economic policy. According to them, only unexpected policy changes lead to changes in real GDP. Robert Lucas explains how people form expectations of the future. He assumes that consumers and firms (economic agents) have accurate information about future economic events. They have rational expectations because they take into account all available information, especially about expected government actions. If the government is following any consistent monetary or fiscal policy, agents know about it and adjust their plans accordingly. So when the government adopts the expected policy measure, it will not be effective because it has been anticipated by agents who have already adjusted their plans. So the government policy will be ineffective.

There is also rational ignorance on the part of agents because they have no incentive to inform themselves about government policy. This ignorance suggests that economists should be cautious when offering policy advice. When policy makers estimate the effect of any policy change, they need to know how people's expectations will respond to the policy change. According to Lucas, using traditional macroeconomic models for policy analysis will be erroneous when used to predict the effects of changes in policy. This is because they do not take into account the impact of policy changes on people's expectations. This is known as the *Lucas critique*.

Let us first take *fiscal policy*. The Keynesians advocate an "activist" fiscal policy to reduce unemployment. But, according to new classicals a tax cut and/or increase in government spending will reduce unemployment only if its short-run effects on the economy are *unexpected* (or unanticipated) by people. In other words, an expansionary fiscal policy may have short-term effects on reducing unemployment provided people do not anticipate that

prices will rise. But when the government persists with such a policy, people expect the rate of inflation to rise. So the workers will press for higher wages in anticipation of more inflation in the future and firms will raise the prices of their products in anticipation of the rise in future costs. As a result, fiscal policy will become ineffective in the short-run. It may cause more unemployment and inflation in the long-run when the government tries to control inflation.

Similarly, if the government adopts an expansionary *monetary policy* by increasing the money supply to reduce unemployment, it is also ineffective in the short-run. Such a policy may reduce unemployment in the short-run provided its effects on the economy are unanticipated. But when the government persists with such an expansionary monetary policy, people expect the inflation rate to rise. Firms raise the prices of their products to overcome the anticipated inflation so that there is no effect on production. Similarly, workers press for higher wages in anticipation of inflation and firms do not employ more workers. So there is no effect on employment.

Thus new classicals suggest that expansionary fiscal and monetary policies will have a temporary effect on unemployment and if continued may cause more inflation and unemployment. For such policies to be successful, they must be unanticipated by the people. Once people anticipate these policies and make adjustments towards them, the economy reverts back to the natural rate of unemployment. Thus for expansionary fiscal and monetary policies to have an impact on unemployment in the short-run, the government must be able to fool the people. But it is unlikely to happen all the time. If the government continues to persist with such policies, they become ineffective because people cannot be fooled for long and they anticipate their effects on production and unemployment. Thus fiscal-monetary policies become ineffective in the short-run. According to new classicals, inflation can be controlled without causing widespread unemployment, if the government announces fiscal and monetary measures and convinces the people about it and do not take them by surprise.

New classicals also restate the monetarist fixed long-run policy rules.

"Since these rules are either ineffective in achieving changes in short-run real output and/or or increase uncertainty, the government should not use them. Instead, it should adopt a constant money supply growth rate rule and a balanced budget."

LAGS IN EFFECTS OF ECONOMIC POLICY

One of the limitations of monetary and fiscal policies in countercyclical manner is the existence of time lags. It takes time for the monetary and fiscal authorities to realise the need for action and its recognition, and the taking of action and the effect of the action on economic activity. According to Friedman, monetary and fiscal actions affect economic conditions only after a lag that is "both long and variable". Friedman distinguishes among three basic lags: the recognition lag, the administrative lag, and the operation lag. These lags are explained as under:

1. The Recognition Lag. It refers to the time between the development of a need for action and the recognition of that need by the monetary and fiscal authorities. It is difficult to know the occurrence of a turning point in a business cycle and recognise the need for action by the authorities. Empirical evidence in the U.S. suggests that in the past the Federal Reserve Bank recognised the need for monetary action only three months after the trough in a business cycle and about six months after a boom had started. Thus the recognition lag has been longer at the peaks than at the troughs.

2. The Administrative Lag. This relates to the period of time that occurs when the monetary and fiscal authorities recognise the need for action and the data on which action is actually taken. The length of the administrative lag (or decision or action lag) varies with the type of policy being considered and the decision-making process of the authorities. Usually, this lag is very short. The administrative lag and the recognition lag taken together are termed *as inside lags* because they fall within the jurisdiction of the authorities. Sometimes, it is difficult to distinguish between the two because the time between recognition of the need for action and the taking of action is so short that the administrative lag becomes the recognition

lag.

3. Operation Lag. The operation lag (or the effects lag) refers to the period of time between the adoption of policy and the final effect of that policy on the economic activity. For analytical convenience, this lag is divided into the intermediate lag and the outside lag.

(a) *The intermediate lag* relates to the moment at which action is taken by the authorities and the moment at which the economy is faced with changes in interest rates, money supply, taxes and public expenditure through monetary and fiscal action.

(b) *The outside lag* refers to the time involved between changes in interest rates, total reserves, credit rationing, etc. and taxes and public spending and their effects on aggregate spending, income and output of the economy.

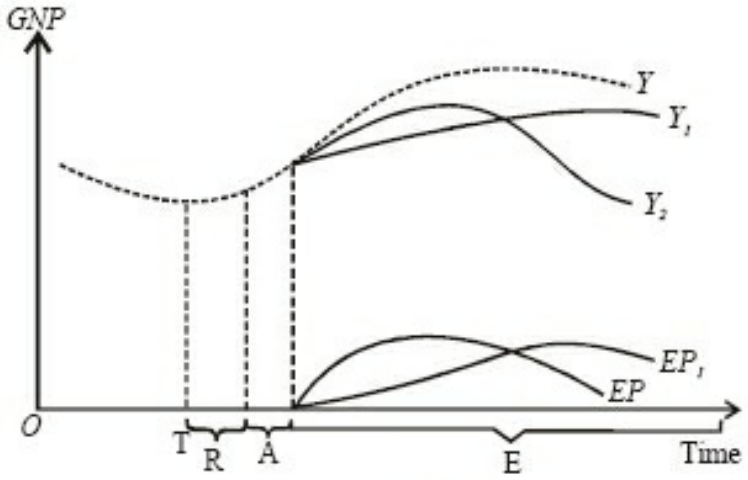


Fig. 5

The three lags are explained in Figure 5 where the time lags are taken on the horizontal axis and aggregate income and output on the vertical axis. Starting from time *T* on the upper turning point of the business cycle, the period *R* shows the recognition lag, *A* the administrative lag and *E* the effect lag. In the effects lag, two alternative effect paths *EP* and EP_1 are shown along with changes in national income as a result of changes in monetary and fiscal policies. The curve *Y* represents the movements in national income before the policy changes. When the effects lag *EP* operates with an expansionary policy to control a downward movement of the business cycle, the curve Y_1 represents the resultant movement in income and output. If the restrictive policy with the effects lag EP_1 is undertaken to control a boom, the resultant path of income is the curve Y_2.

According to Friedman, a lag is both long and variable and existence of long and variable lags in the effects of changes has led him to conclude that discretionary countercyclical policy aimed at economic stabilisation may be destabilising. These lags can intensify rather than mitigate cyclical fluctuations. Let us explain Friedman's view in detail.

Countercyclical policy means "leaning against the prevailing economic winds". It implies that the authorities follow an easy policy in a recession and a tight policy in a boom. Given the long length of the lags and their variability, the effects of an easy monetary and fiscal policy to control a recession may lead to inflationary pressures. Similarly, a tight monetary and fiscal policy to control a boom may lead to recessionary conditions due to time lags and their delayed effects. According to Friedman, "We seldom in fact know which way the economic wind is blowing until several months after the event." Due to unpredictability in forecasting of booms and recessions and the length, variability and uncertainty of time lags, a countercyclical policy is destabilising rather than stabilising. Since the implementation and effects of anticyclical policy do not occur at the right time, they will add to the amplitude of the cycle and destabilise the economy. Friedman, therefore, favours an "automatic framework" of monetary and fiscal policy in which the money supply and fiscal deficit increase at a steady and inflexible rate. He gives two reasons in support of the automatic framework as against the discretionary policy action. *First*, discretionary anticyclical policy is often dominated by goals other than, and even contradictory to, stabilisation when the monetary and fiscal authorities adopt such measures as pegging bond yields, halting gold outflows, etc. But the automatic framework cannot be easily exploited for other purposes. *Second*, the automatic framework would be free from inertia and political considerations that inhibit the reversal of discretionary policies when they turn out to be in the wrong direction.

Friedman, therefore, calls for an end to discretion in policy. In place of the judgement of monetary and fiscal authorities, he proposes that it should follow a fixed long-run rule, that is to increase the money supply and fiscal deficit at an annual fixed percentage rate regardless of current economic conditions. Set the money supply and fiscal deficit at a fixed rate and leave it to grow at that rate automatically. Of course, such an

automatic framework would not eliminate economic fluctuations, but it would at least prevent the economy from the perverse effects of delayed discretionary action. For instance, Friedman is not rigid about any exact rate of increase in the money supply. He suggests a rate in the neighbourhood of 3 to 4 per cent per year which closely approximates the economy's past *average* annual rate of growth in real output. Such a stable growth rate of money supply will keep prices stable and allow the aggregate demand to grow with the growth in productive capacity of the economy. It will thus increase employment and income. Any inflationary situation will be temporary because it will not receive an impetus by any increase in the money supply. Similarly, any recessionary tendency will be short-lived because the liquidity provided by the constant money supply will cause aggregate demand to expand. Thus the automatic framework suggested by Friedman would remove uncertainties of forecasting economic fluctuations and the problems of time lags associated with discretionary countercyclical monetary policy. Figure 6 illustrates the nature of business cycles in the case of two types of policies. The dotted curve shows the fluctuations in economic activity when countercyclical monetary policy is adopted. The smooth curve shows mild fluctuations in economic activity under the automatic framework.

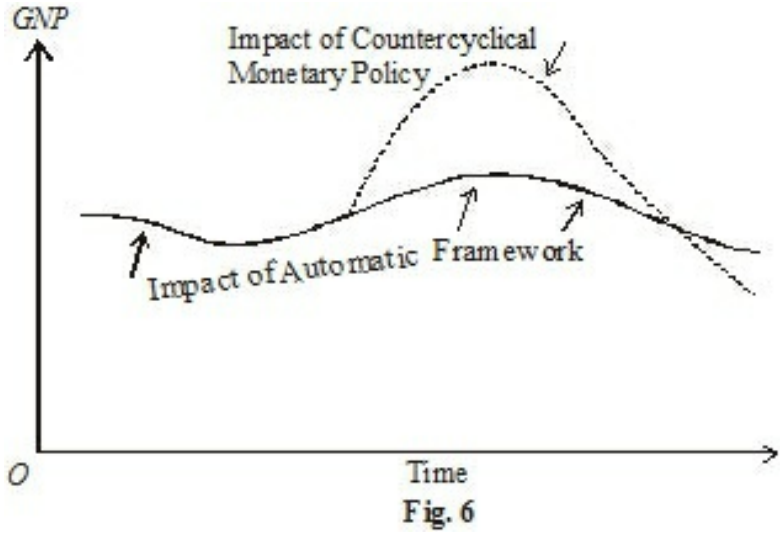


Fig. 6

Its Criticisms

Culbertson does not agree with Friedman's automatic policy framework to eliminate unpredictability and lag problems of discretionary policy. He opines that "the broad record of experience supports the view that anticyclical monetary, debt management and fiscal adjustments can be counted on to have their prominent direct effects within three to six

months, soon enough that if they are undertaken moderately early in a cyclical phase they will not be destabilising".

The Keynesians also do not concur with Friedman's policy prescription to avoid the lag problems. They favour the use of monetary policy to control a boom, and supplementing monetary policy with fiscal policy to control a recession.

Whichever policy is adopted, the lag problems and economic fluctuations cannot be completely eliminated.

SELECTED READINGS

A.K. Dasgupta and A.J. Hagger, *The Objectives of Macroeconomic Policy*, 1977.

G.K. Shaw, *An Introduction to the Theory of Macroeconomic Policy*, 1971.

N.F. Kaiser, *Readings in Macroeconomics*, 1970.

R.E. Hall and J.B. Taylor, *Macroeconomics*, 5/e, 1997

R.T. Froyen, *Macroeconomics*, 7/e, 2002.

EXERCISES

1. Discuss the principal objectives of macroeconomic policy.
2. Is there any conflict between the objectives of full employment and price stability ? If so, how will you reconcile them ?
3. Discuss some of the problems of economic policy. Explain particularly the trade-off in policy objectives.
4. Indicate the way and the extent to which monetary and fiscal instruments could be successfully used to bring about internal and external balance.
5. What policy instruments would you suggest for maintaining both internal and external equilibrium ?

6. “Fiscal policy should always be directed toward the achievement of internal balance. This is most obviously true when a policy ‘conflict’ arises.” Discuss.
7. Explain Swan’s model of internal and external balance.
8. Discuss Mundell’s model of maintaining external and internal balance.
9. What is assignment problem ? How has Mundell solved it ?
10. Analyse the controvesy over rules vs. discretion in economic policy.
11. Explain the effects of lags in economic policy.

10. Write short notes on : Okun’s Law, Tinbergen Principle, Expectations in Economic Policy.

CHAPTER 46

MONETARY POLICY

MEANING OF MONETARY POLICY

Monetary policy refers to the credit control measures adopted by the central bank of a country. Johnson defines monetary policy “as policy employing central bank’s control of the supply of money as an instrument for achieving the objectives of general economic policy.” G.K. Shaw defines it as "any conscious action undertaken by the monetary authorities to change the quantity, availability or cost of money."

OBJECTIVES OR GOALS OF MONETARY POLICY

The following are the principal objectives of monetary policy:

1. Full Employment

Full employment has been ranked among the foremost objectives of monetary policy. It is an important goal not only because unemployment leads to wastage of potential output, but also because of the loss of social standing and self-respect.

2. Price Stability

One of the policy objectives of monetary policy is to stabilise the price level. Both economists and laymen favour this policy because fluctuations in prices bring uncertainty and instability to the economy.

3. Economic Growth

One of the most important objectives of monetary policy in recent years has been the rapid economic growth of an economy. Economic growth is defined as "the process whereby the real per capita income of a country increases over a long period of time."

4. Balance of Payments

Another objective of monetary policy since the 1950s has been to maintain equilibrium in the balance of payments.

INSTRUMENTS OF MONETARY POLICY

The instruments of monetary policy are of two types: *first*, quantitative, general or indirect; and *second*, qualitative, selective or direct. They affect the level of aggregate demand through the supply of money, cost of money and availability of credit. Of the two types of instruments, the first category includes bank rate variations, open market operations and changing reserve requirements. They are meant to regulate the overall level of credit in the economy through commercial banks. The selective credit controls aim at controlling specific types of credit. They include changing margin requirements and regulation of consumer credit. We discuss them as under:

Bank Rate Policy. The bank rate is the minimum lending rate of the central bank at which it rediscounts first class bills of exchange and government securities held by the commercial banks. When the central bank finds that inflationary pressures have started emerging within the economy, it raises the bank rate. Borrowing from the central bank becomes costly and commercial banks borrow less from it. The commercial banks, in turn, raise their lending rates to the busniess community and borrowers borrow less from the commercial banks. There is contraction of credit and prices are checked from rising further. On the contrary, when prices are depressed, the central bank lowers the bank rate. It is cheap to borrow from the central bank on the part of commercial banks. The latter also lower their lending rates. Businessmen are

encouraged to borrow more. Investment is encouraged. Output, employment, income and demand start rising and the downward movement of prices is checked.

Open Market Operations. Open market operations refer to sale and purchase of securities in the money market by the central bank. When prices are rising and there is need to control them, the central bank sells securities. The reserves of commercial banks are reduced and they are not in a position to lend more to the business community. Further investment is discouraged and the rise in prices is checked. Contrariwise, when recessionary forces start in the economy, the central bank buys securities. The reserves of commercial banks are raised. They lend more. Investment, output, employment, income and demand rise and fall in price is checked.

Changes in Reserve Ratios. This weapon was suggested by Keynes in his *Treatise on Money* and the USA was the first to adopt it as a monetary device. Every bank is required by law to keep a certain percentage of its total deposits in the form of a reserve fund in its vaults and also a certain percentage with the central bank. When prices are rising, the central bank raises the reserve ratio. Banks are required to keep more with the central bank. Their reserves are reduced and they lend less. The volume of investment, output and employment are adversely affected. In the opposite case, when the reserve ratio is lowered, the reserves of commercial banks are raised. They lend more and the economic activity is favourably affected.

Selective Credit Controls. Selective credit controls are used to influence specific types of credit for particular purposes. They usually take the form of changing margin requirements to control speculative activities within the economy. When there is brisk speculative activity in the economy or in particular sectors in certain commodities, and prices start rising, the central bank raises the margin requirement on them. The result is that the borrowers are given less money in loans against specified securities. For instance, raising the margin requirement to 60% means that the pledger of securities of the value of Rs 10,000 will be given 40% of their value, i.e. Rs 4,000 as loan. In case of recession in a particular sector, the central bank encourages borrowing by lowering margin requirements.

Conclusion. For an effective anticyclical monetary policy, bank rate, open market operations, reserve ratio and selective control measures are required to be adopted simultaneously. But it has been accepted by all monetary theorists that (*i*) the success of monetary policy is nil in a depression when business confidence is at its lowest ebb; and (*ii*) it is successful against inflation. The monetarists contend that as against fiscal policy, monetary policy possesses greater flexibility and it can be implemented rapidly.[1]

EXPANSIONARY MONETARY POLICY

An expansionary (or *easy*) monetary policy is used to overcome a recession or a depression or a deflationary gap. When there is a fall in consumer demand for goods and services, and in business demand for investment goods, a deflationary gap emerges. The central bank starts an expansionary monetary policy that eases the credit market conditions and leads to an upward shift in aggregate demand. For this purpose, the central bank purchases government securities in the open market, lowers the reserve requirements of member banks, lowers the discount rate and encourages consumer and busniess credit through selective credit measures. By such measures, it decreases the cost and availability of credit in the money market, and improves the economy.

1. For limitations of Monetary Policy, refer to the Scope and Limitations given below under Expansionary and Restrictive Monetary Policy.

The expansionary (*cheap*) monetary policy is explained in terms of Figure 1 (A) and (B) where the initial recession equilibrium is at *R*, *Y*, *P* and *Q*. At the interest rate *R* in Panel (A) of the figure, there is already an excess money supply in the economy. Suppose the central bank credit policy results in an increase in the money supply in the economy. This leads to a rightwards shift of the *LM* curve to LM_1. This increases income from *OY* to OY_1 and aggregate demand expands and the demand curve *D* shifts upwards to D_1 in panel (B). With the increase in the demand for goods and services, output increases from *OQ* to OQ_1 at a higher price level P_1. If the expansionary monetary policy operates smoothly, the equilibrium at

E_1 can be at the full employment level. But this is not likely to be attained because of the following limitations.

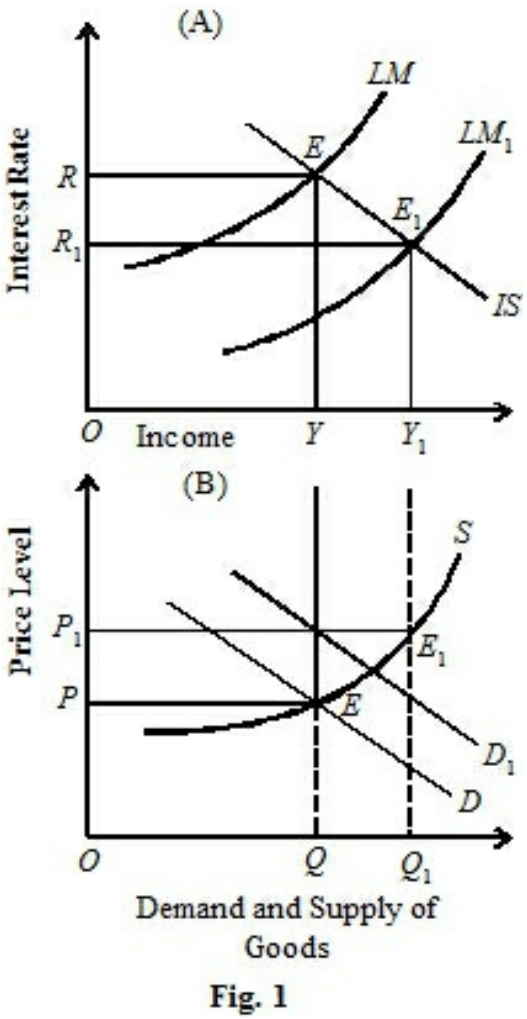


Fig. 1

Its Scope and Limitations

During the 1930s and 1940s, it was believed that the success of monetary policy in stimulating recovery from a depression was severely limited than in controlling a boom and inflation. This view emerged from the experiences of the Great Depression and the appearance of Keynes's *General Theory*.

The monetarists hold that during a depression the central bank can increase the reserves of commercial banks through a cheap money policy. They can do so by buying securities and reducing the interest rate. As a result, their ability to extend credit facilities to borrowers increases. But the experience of the Great Depression tells us that in a serious depression when there is pessimism among businessmen, the success of such a policy is practically nil. In such a situation, banks are helpless in bringing about a revival. Since business activity is almost at a stand still, businessmen do not have any inclination to borrow to build up inventories even when the rate of interest is very low. Rather, they want to reduce their inventories by repaying loans already drawn from the banks. Moreover, the question of borrowing for long-term capital needs does not arise in a depression when the business activity is already at a very low level. The same is the case with consumers who faced with unemployment and reduced incomes do not like to purchase any durable goods through bank loans. Thus all that the banks can do is to make credit available but they cannot force businessmen and consumers to accept it. In the 1930s, very low interest rates and the piling up of unused reserves with the banks did not have any

significant impact on the depressed economies of the world.

“This is not to say that an easy monetary policy in times of severe contraction will be without beneficial effect, its effect will be largely that of preventing a bad situation from getting worse. *But* a restrictive monetary policy combined with a business downturn would surely aggravate the downturn—the classical example of this was the monetary policy in 1931 that contributed to the deepening of the Great Depression...On the other hand, if credit is readily available on favourable terms, it clearly has a stabilising effect. By meeting the liquidity requirments of business, it can slow and perhaps reduce the extent of the downturn."[2]

But what led to the decline of monetary policy in the 1930s and 1940s ? In addition to the sad and disillusioning experiences during and after the Great Depression, it was Keynes’s *General Theory* that led to a decline in monetary policy as an instrument of economic stabilisation. Keynes pointed out that a highly elastic liquidity preference schedule (*liquidity trap*) renders monetary policy impotent in time of severe depression.

RESTRICTIVE MONETARY POLICY

A monetary policy designed to curtail aggregate demand is called restrictive (or *dear*) monetary policy. It is used to overcome an inflationary gap. The economy experiences inflationary pressures due to rising consumers’ demand for goods and services and there is also boom in business investment. The central bank starts a restrictive monetary policy in order to lower aggregate consumption and investment by increasing the cost and availability of bank credit. It might do so by selling government securities in the open market, by raising reserve requirements of member banks, by raising the discount rate, and controlling consumer and business credit through selective measures. By such measures, the central bank increases the cost and availability of credit in the money market and thereby controls inflationary pressures.

The restrictive monetary policy is also explained in terms of Figure 1 (A) and (B) where the initial recession equilibrium is at R_1, Y_1, P_1 and Q_1. At

the interest rate R_1 in Panel (A) of the figure, there is already an excess money supply in the economy. Suppose the central bank credit policy results in a decrease in the money supply in the economy. This leads to a leftward shift of the LM_1 curve to *LM*. This decreases income from OY_1 to *OY* and aggregate demand falls and the demand curve D_1 shifts downwards to *D* in panel (B). With the decline in the demand for goods and services, output falls from OQ_1 to *OQ* at a lower price level *P*.

Its Scope and Limitations

But the scope of monetary policy is severly limited in controlling inflation. The following are its *limitations*.

1. Increase in the Velocity of Money. One of the important limitations on the effectiveness of monetary policy in controlling inflation is the increase in the velocity of money held by the public. The central bank can control the money supply and the cost of money by a tight monetary policy but it does not possess any power to control the velocity of money. The public can make an effective use of the money supply held by them thereby making a restrictive monetary policy ineffective. This can be done in a number of ways.

(a) Commercial Bank Portfolio Adjustments. In the face of a restrictive monetary policy, commercial banks meet the borrowers' demand for loans by selling government securities to the central bank. Such a policy simply converts idle deposits held by the banks in the form of securities into active deposits. Government securities lying in the bank's portfolios are substituted for loans. But there is no change in either the total deposits or the money supply with the banks. However, this leads to increase in total spending when the banks lend money to borrowers. Thus the restrictive monetary policy of the central bank becomes ineffective.

---

2. Edward Shapiro, *op. cit*., p.541. Italics mine.

Moreover, when the banks sell government securities to the central bank, their prices fall and the interest on them rises in the market. This will raise the general interest rate structure in the market. But the fall in the prices of

securities brings capital losses to the banks and they may be reluctant to bear them. This depends upon whether they expect the fall in security prices (or rise in interest rate) to be short-lived or continue overtime. If the fall in security prices is expected to be short-lived, the banks will prefer to keep securities rather than sell them at a capital loss. On the other hand, if they expect it to continue for some time, they will sell securities for giving loans to customers at higher interest rates, thereby recouping the capital loss on the sale of securities through higher interest rates on loans. But once the demand for loans subsides, the banks can buy back government securities now at prices lower than at which they sold, and again gain in the transaction. Thus the commercial banks' policy of portfolio adjustment raises the velocity of total money supply even in the face of a tight monetary policy thereby making the latter ineffective.

(b) The Role of Non-Bank Financial Intermediaries. NBFIs act as a restraint on the effectiveness of monetary policy to restrict the money supply in two ways. *First*, they sell securities for advancing loans, and thus increase velocity in the same manner as commercial banks do, as explained above. *Second*, as interest rates on securities rise in a tight monetary policy, financial intermediaries raise the interest rates on deposits with them to attract more funds from savers. This induces savers to shift more idle money to the intermediaries which increases their lending power further. In this way, they are able to raise the velocity of money thereby making tight restrictive monetary policy ineffective.

(c) Methods to Make Better Use of Available Money Supply. The private sector has evolved many ways to make better use of available supply of money which make a restrictive monetary policy ineffective. Some of the methods are the evolvement of improved methods of collecting funds by sales finance companies, borrowing funds by companies from the public at higher rates than offered by commercial banks, etc. By getting funds from sources other than the commercial banks, such institutions are able to increase the velocity of the available supply of money even under restrictive monetary policy.

2. Discriminatory. A restrictive monetary policy is discriminatory in its effects on particular sectors of the economy. It is argued that firms that

depend upon internal source of financing are not affected by a restrictive monetary policy. On the other hand, only those firms are affected that depend for funds on the banking system. In particular, a tight monetary policy "*is thought to work against small businessmen, becuase they are poorer credit risks, and against residential construction and some types of state and local government spending, because they are most sensitive to changes in credit cost.*" It may slow down or even halt spending by them.

3. Threat to Credit Market. If the central bank rigorously tightens the credit market and investors expect continued increases in interest rates, this may lead to the drying up of loanable funds to the credit market. As a result, securities may not be sold and the credit market may cease to function.

4. Threatens Solvency of NBFIs. A vigorous monetary policy by swiftly raising interest rates may threaten the solvency of such NBFIs as savings banks, and savings and loan associations. This is because unlike the commercial banks, they are not in a position to adjust themselves to rapidly increasing interest rates.

5. Alter Expectations of Borrowers and Lenders. A very tight monetary policy may alter the expectations of borrowers and lenders. So they bring irreversible changes in credit market conditions. A rapid rise in interest rates may so change expectations that even when this policy is abandoned and an expansionary policy is started, lenders may be reluctant to make long-term loans in anticipation of rise in interest rates again. On the other hand, borrowers may borrow long-term funds even if they do not need them immediately in anticipation of rise in interest rates in the future.

6. Time Lags. Another important limitation of a tight monetary policy is the existence of time lags which are related to the need of action, its recognition, and the decision and operation of actions in time. As the monetary authority is not able to adopt restrictive monetary measures in time due to these time lags, monetary policy works very slowly and hence it is not very effective in controlling inflation.

EXERCISES

1. Explain the principal instruments of monetary policy. What are their limitations ?
2. Critically discuss the effectiveness of monetary policy in controlling inflation.
3. Distinguish between cheap and dear monetary policy. Explain their relative effectiveness.

CHAPTER 47

CLASSICAL, KEYNESIAN AND MODERN VIEWS ON MONETARY POLICY

THE CLASSICAL VIEW

Money, according to the classicists, is a veil. It is neutral in its effects on the economy. It simply affects the price level, but nothing else. An increase in the money supply leads to an increase in the price level, but the real income, the rate of interest and the level of real economic activity remain unaffected.

In the classical system, the main function of money is to act as a medium of exchange. It is to determine the general level of prices at which goods and services will be exchanged. This relationship between money and the price level is explained in terms of the quantity theory of money. The classical quantity theory of money states that the price level is a function of the supply of money. Algebraically, $MV=PT$ where M, V, P, and T are the supply of money, velocity of money, price level and the volume of transactions (or real total output). The equation tells that the total money supply MV equals the total value of output PT in the economy. Assuming V (the velocity of money) and T (the total output) to be constant, a change in the supply of money (M) causes a proportional change in the price level (P).

The classicists believed that there was always full employment in the economy. At the same time, they recognised the existence of unemployment in the event of downward rigidity of money wages. Such a situation could be corrected by an expansionary monetary policy. The

process involved is as follows. Suppose the monetary authority increases the money supply, given the velocity of money and the level of real output. With the increase in the money supply, liquidity rises with the people who increase the demand for goods and services. This, in turn, raises the price level. The rise in price level reduces the real wage which provides incentive for employers to expand employment and output toward the full employment level. Thus an expansionary monetary policy is effective in restoring full employment in the classical system.

Keynes did not agree with the classical view that the supply of money influences the price level directly and that the economy always stays at the full employment level. Moreover, the classical analysis was related to the long-run where market forces worked the economy toward full employment.

THE KEYNESIAN VIEW

In the Keynesian analysis, monetary policy plays a crucial role in affecting economic activity. It contends that a change in the supply of money can permanently change such variables as the rate of interest, the aggregate demand, and the level of employment, output and income.

Keynes believed in the existence of unemployment equilibrium. This implies that an increase in money supply can bring about permanent increases in the level of output. The ultimate influence of money supply on the price level depends upon its influence on aggregate demand and the elasticity of the supply of aggregate output.

In a situation of unemployment, Keynes advocated cheap money policy. When the supply of money is increased, its first effect is on the rate of interest which tends to fall. Given the marginal efficiency of capital, a fall in the rate of interest will increase investment. The increased investment will raise effective demand through the multiplier effect thereby increasing income, output and employment. Since the supply curve of factors of production is perfectly elastic in a situation of unemployment, wage and non-wage workers are available at a constant rate of remuneration. There being constant returns to scale, prices do not rise with

the increase in output so long as there is unemployment. Under the circumstances, output and employment will increase in the same proportion as effective demand, and the effective demand will increase in the same proportion as the supply of money. But once full employment is reached, output ceases to respond at all to changes in the supply of money and so in effective demand. An increase in the supply of money beyond the level of full employment would raise the price level in the classical fashion.

What causes the rate of interest to change in the Keynesian monetary policy? In the Keynesian analysis, the rate of interest is determined by the demand for and supply of money. If either the demand for money or the supply of money changes, the equilibrium rate of interest would change. The supply of money is determined by the monetary authority which is normally fixed in the short-run. In other words, the money supply curve is perfectly inelastic. The demand for money is the desire to hold cash for transactions, precautionary and speculative purposes. Money held for transactions and precautionary motives depends upon the level of income. The speculative demand for money depends upon the rate of interest or bond prices.

It is, in fact, expectations about changes in bond prices or in the market rate of interest that determine the speculative demand for money. Keynes considered only two types of assets in his analysis, money and bonds. Money does not yield anything explicitly and bonds pay an explicit rate of interest. Therefore, people wish to hold bonds rather than liquid money because interest is paid to bondholders. Keynes held that the demand for money is a decreasing function of the rate of interest. The higher the rate of interest, the lower the demand for money, and vice versa. This negative relationship between the demand for money and the rate of interest provides a link between changes in the supply of money and the level of economic activity.

Keynes himself proceeded to question the efficacy of his monetary policy under certain conditions. He argued that at a very low interest rate, the demand for money curve becomes perfectly elastic. This is the *liquidity trap* portion of the demand for money curve which is completely flat. This

means that further increases in the supply of money by the monetary authority cannot reduce the rate of interest. This implies that there will be no effect on investment and income, and monetary policy does not influence economic activity. Given an interest-inelastic investment function, monetary policy will be ineffective in the Keynesian analysis.

Thus Keynes believed on the basis of his experience that monetary policy operated under certain limitations. He wrote, “If we are tempted to assert that money is the drink which stimulates the system of activity, we must remind ourselves that there may be several steps between the cup and the lip.” The effectiveness of monetary policy depends on: *first,* if the increase in the supply of money reduces the rate of interest provided the demand for money does not become infinite (i.e. perfectly elastic), and *second*, the reduction in the rate of interest increases investment demand provided it is not inelastic to the rate of interest. These limitations become more serious during depression and thus monetary policy becomes ineffective. That is why, Keynes favoured investment on public projects during depression. He wrote, “I am now somewhat sceptical of the success of a merely monetary policy directed towards influencing the rate of interest.” In fact, he advocated supplementing monetary policy with fiscal policy during depression. Recent researches have shown that Keynes was misrepresented by his followers in attributing that he was not a votary of monetary policy.

THE MODERN VIEW

The modern monetary economists[1] reject the Keynesian view that the link between the supply of money and output is the rate of interest. The Keynesian analysis considered only two types of assets: bonds and speculative cash balances, and their allocation depended on the rate of interest which, in turn, resulted in changes in output.

The modern monetary policy is based on the portfolio adjustment process. When the central bank purchases securities in the open market, it sets in motion substitution and wealth effects, as the public portfolio consists of a wide variety of assets such as bonds, equities, savings, mortgages, etc. These effects will ultimately increase aggregate money demand and

expand output. The transmission mechanism is explained as under.

---

1. For a detailed discussion of the monetarist view refer to chapter *The Monetarist Revolution.*

Substitution Effects

The neo-Keynesians widened considerably the portfolio of assets to include not only government securities but also industrial bonds, equities, savings, mortgages, etc. Given this type of portfolio, suppose the central bank engages in open market operation and purchases securities. This will increase the prices of securities, thereby reducing the yield on them. In other words, the holders of securities sell them to the central bank because they get high prices for them. They now hold more money than they desire. As a result, they try to readjust the structure of their portfolios so as to reduce their money holdings. Suppose they substitute bonds for their excess money balances. The increase in the demand for bonds results in an increase in their market price, thereby reducing their current yield, as interest rate falls. Consequently, the demand for other assets such as equities, consumer durables etc. increases. When people having surplus money balances purchase equities (shares), their prices rise. As a result, the value of capital of such firms rises above the supply price of such new capital. Such firms are, therefore, induced to increase their demand for more capital equipment, thereby raising output in the capital goods industries. This will, in turn, spread to the rest of the economy via the multiplier effect. Thus the "neo-Keynesians contend that financial assets are the closest substitutes for money, and that, consequently, increases in the supply of money will have their effect eventually on the level of economic activity by bringing about increases in the output of capital goods industries."

The monetarists led by Friedman are of the view that excess money balances will be used to purchase not only financial assets but also real assets such as houses, land, consumer durables, etc. So when the central back purchases securities and their prices increase and yields fall, the demand for financial and real assets increases. With the increase in their demand, their prices rise. But the rise in the prices of real assets leads to some additional effects. When the prices of real assets rise, their

production is encouraged which, in turn, raises the demand for resources used in their production. Moreover, there is also increase in the demand for services as a result of increase in the prices of real assets. To illustrate, suppose the holders of surplus money balances demand more consumer durables, say cars. This will raise the prices of existing cars relative to the prices of new cars. Since the new cars are relatively cheaper, their demand will increase which will raise their output, income and employment in the car industry. With the overall increase in the demand for cars, the demand for their services will also increase. Thus an expansionary monetary policy leads to increase in demand, prices and expenditures for financial and real assets and for services through substitution effect.

The transmission mechanism operates through the initial change in interest rates on securities and relative prices of both financial and real assets. These changes lead to the substitution of asset holdings which imply changes in the demand for real assets and services.

Wealth Effects

There is a lot of controversy between the neo-Keynesians and the Friedmanians regarding the wealth effect. According to the neo-Keynesians when the central bank engages in open market purchasing of securities, it leads to decline in market interest rates. This produces a wealth effect which results from the fact that at a lower rate of interest the present or the capitalised value of the expected income stream of financial or real capital assets increases. As people feel wealthier, they buy more of all assets in their portfolios, and consequently increase their demand for consumer nondurables.

According to the Friedmanians, an open market operation which exchanges money for bonds increases the nominal wealth of society. This is a *direct* wealth effect. A fall in the rate of interest increases the market value of the existing capital stock, thereby raising the nominal wealth of society. This is the *interest-induced* wealth effect. Both these wealth effects lead to an increase in net wealth. As a result, individuals buy financial and real assets including consumer durables.

In fact, the extent to which “the combined wealth effects bring up the

price of the existing stock of assets relative to the price of new production, they also serve to increase the rate of flow of output of new producer and consumer durables. In addition, the wealth effect may serve directly to stimulate the purchase of consumer nondurable goods."

EXERCISES

1. Discuss the Classical and Keynesian views on monetary policy.
2. Explain the modern view on monetary policy.

CHAPTER 48

THE LIQUIDITY THEORY OF MONEY

INTRODUCTION

The liquidity theory is a new approach to monetary policy which has been propounded by the Radcliffe Report and Gurley and Shaw. This theory highlights the role of liquid assets in influencing aggregate demand and economic activity. According to this theory, it is non-bank financial intermediaries (NBFIs) which provide a link between liquidity and monetary policy.

Broadly speaking, liquidity means *moneyness.* Liquidity is characterised by the ease of converting an asset into money at little cost. A liquid asset is one which is easily spendable, marketable (transferable) and has capital certainty.

Depending on these properties, different assets differ in the degree of liquidity. Money is the most important liquid asset which consists of coins, currency and bank deposits. This is because they perform the medium of exchange function of money. In the case of bank deposits only *demand deposits* are money because cheques can be issued on their basis. Thus cheques are perfect substitutes for money. But time deposits are *not real* money but are *near money* because they cannot be *withdrawn before* the expiry of a fixed period.

Besides time deposits, other near money assets are bonds, securities, debentures, bills of exchange, treasury bills, insurance policies, etc. All these types of assets are *less* liquid because they involve cost for converting them into real money.

Liquidity is created by the central bank by issuing coins and notes and by commercial banks by creating demand deposits. A change in the supply of money by the central bank affects liquidity by bringing changes or readjustments in the portfolio assets of the people. This depends on the effect of money supply on aggregate spending. For instance, if people decide to spend the increased money supply in purchasing such assets as shares and debentures, there will be less money available in liquid form with the public. If the stock market is bullish, people may buy shares and the liquidity is reduced. On the other hand, if there is uncertainty in the stock market, people may hold the increased money supply in bank deposits or invest in property if they expect property prices to rise. But it is the central bank that influences money supply in the economy by following "easy" or "tight" monetary policy. It does so by controlling aggregate spending and thereby influencing business activity, output and employment. But it is not always successful in controlling spending by increasing or decreasing the money supply and hence liquidity. This is because the central bank has little control over the velocity of circulation of money, non-bank financial intermediaries, business expectations, time lags in monetary policy, etc. It is, therefore, very difficult to predict the effects of changes in money supply on liquidity.

. The Radcliffe Report and Gurley and Shaw have shown separately that the existence of NBFIs weaken the effect of monetary policy in increasing the money supply in the economy because they control a large portion of liquidity in the economy. Therefore, the role of NBFIs in increasing liquidity is equally important. We explain their views as under.

THE RADCLIFFE COMMITTEE VIEW : RADCLIFFE - SAYERS THESIS

The Radcliffe Committee in its report did not accept the monetarist view that there was any direct link between money supply and national income. This is because money is a close substitute for other financial assets, particularly those issued by non-bank financial institutions. If the central bank wanted to follow a contractional monetary policy by selling securities in the open market, non-bank financial intermediaries would be able to release idle demand deposits and currency. In this way, they would offset the restriction of money supply, and leave aggregate demand

unchanged. Thus it held the view that “changes in rates of interest only very exceptionally have direct effects on the level of demand.” It also did not find any indirect relation between money supply and national income.

According to the Radcliffe Committee, it is liquidity that influences total effective demand for goods and services rather than the supply of money The Committee pointed out that the “decisions to spend on goods and services, the decisions that determine the level of total demand, are influenced by the liquidity of the spenders. The spending is not limited by the amount of money in existence; but it is related to the amount of money people think they can get hold of, whether by receipt of income (for instance, from sales), by disposal of capital assets or by borrowing”. Thus liquidity consists of the amount of money people can get hold of from income, sales, disposal of capital assets or borrowing. Suppose if liquidity is reduced, expenditures would decline to the extent they exceed current income and vice versa.

The second element by which the monetary authority can influence the level of overall liquidity and hence the level of aggregate expenditures is the *interest rate*. It is by manipulating the structure of interest rates which affects the overall liquidity structure. A movement of interest rates means significant changes in the capital values of many assets held by individuals and financial institutions. When interest rates rise, they reduce liquidity because they reduce the capital value of assets held by the latter. Financial institutions, in turn, reduce the supply of loanable funds. Since individuals and businesses cannot procure funds, they reduce their expenditures. On the other hand, a fall in the interest rate “strengthens balance sheets and encourages lenders to seek new business”. In this monetary mechanism, the commercial banks hold a special position in the Committee’s view, because they are the most convenient institutional source of funds for most borrowers for short-term purposes.

On the basis of this transmission mechanism, the Committee concluded that monetary policy performed a “background function”. It thought that business investment was insensitive to changes in short-term interest rates. But it was influenced by long-term interest rates. Therefore, it was undesirable to cause long-term interest rates to fluctuate sharply because

sharp fluctuations would undermine the strength and stability of financial institutions. In fact, financial institutions are accustomed to stable interest rates and they plan their lending policies accordingly. The monetary authority should, therefore, try to hold long-term interest rates at the level which maintained approximate balance between saving and investment in the economy.

The Committee did not favour reliance on *monetary policy* alone in severe deflationary or severe inflationary conditions. Apart from the removal of credit control and a possible effect of lower long-term interest rates on house-building activity during a slump, monetary policy was thought to be of not much help. Rather, the use of monetary measures to raise the economy from a slump carried the danger of flooding the financial system with liquidity. It would be difficult to control it later on when business activity expanded during the upswing. Thus in the Committee's view, "in general, the potentialities of monetary policy alone in the face of a severe slump are well represented by the proverb that you can take a horse to water but you cannot make him drink." Similarly, it was against any restriction on the money supply during severe inflation. It advocated measures to strike directly and rapidly at the liquidity of spenders through control of capital issues, bank advances and consumer credit. But it did not favour control over the lending power of the non-bank financial intermediaries because it entailed additional administrative burdens. More so because new financial institutions would be coming up which would make it difficult for the monetary authority to control the situation. Thus the Radcliffe Committee suggested "liquidity controls" through changes in long-term interest rates. According to it, "monetary measures cannot alone be relied upon to keep in nice balance an economy subject to major strains from without and within. Monetary measures can help but that is all". It, therefore, recommended reliance on fiscal policy in normal times which should be supplemented with monetary policy during severe inflations.

Its Criticisms

The views of the Radcliffe Committee have been severely criticised on the following grounds:

1. Economists have criticised it for advocating the control of overall liquidity of the economy rather than the money supply.

2. It did not precisely define liquidity or liquid assets. This makes the whole approach to liquidity as haphazard.

3. Prof. Gurley in his review of the Radcliffe Report found "confusion every where in the role of money supply, in the concept of liquidity, and in money to satisfy the liquidity desire of the public."

4. Other economists characterised the Report as "an utterly muddled one, indicating a stock-flow muddle, a liquidity muddle, and an interest rate muddle."

THE GURLEY-SHAW VIEW

According to Gurley and Shaw, it is the NBFIs that provide liquidity and safety to financial assets and help in transferring funds from ultimate lenders to ultimate borrowers for productive purposes. They increase capital formation and consequently lead to economic growth.

By buying primary securities from the ultimate borrowers and selling indirect securities to the ultimate lenders, the intermediaries influence the availability of credit and the structure and level of interest rates. They create credit different from commercial banks. But they create new assets and liabilities which tend to influence the supply of money and thus hinder the operation of an effective monetary policy.

According to them, the savings deposits of NBFIs resemble the demand deposits of commercial banks because it is not difficult for NBFIs to convert their savings deposits into cash. These savings deposits, whether of commercial banks or of NBFIs, are for all practical purposes as liquid as demand deposits. Such savings deposits held by NBFIs are known as near moneys. Since the demand deposits are not controlled by the central bank, it follows that the savings deposits held by NBFIs will hinder the successful operation of a successful monetary policy.

If the central bank wishes to control excess liquidity in the economy only through the reduction in money supply, it will not be successful because the savings deposits of NBFIs can be converted into cash. Similarly, if the central bank tries to control lending by commercial banks, it will not be successful if the lending of all other FIs are not under its control, as is the case. According to Gurley and Shaw, this problem arises especially when the central bank adopts an anti-inflationary monetary policy. Suppose the central bank reduces the money supply in order to control inflation. Among other effects, the interest rates on market securities rise in anticipation of higher yields and profits. NBFIs will raise the interest rates on their savings deposits to attract more funds in order to invest them in higher yielding securities.

Persons already holding securities find that their prices have fallen because of the rise in interest rates on present securities. They will, therefore, sell them and deposit their funds with intermediaries in order to earn higher interest rates on savings deposits. In the meantime, attracted by higher interest rates, others holding idle cash balances will also deposit them with intermediaries. So when NBFIs raise the interest rates on their savings deposits, the public reduces its demand for money which, in turn, reduces the market rate of interest. Thus NBFIs make tight monetary policy less successful or effective. Similarly, NBFIs can make an expansionary monetary policy ineffective by reducing liquidity.

But unlike the Radcliffe Report, Gurley and Shaw argue that the central bank's control over NBFIs should be extended for an effective monetary policy. This is because the NBFIs create more near-money assets and thereby affect the overall liquidity which, in turn, influences aggregate demand and economic activity.

Its Criticisms

The view of Gurley and Shaw has been criticised on the following grounds:

1. Prof. Johnson does not agree with Gurley and Shaw. He observes that there seems to be no empirical case for empowering the central bank to extend its control over financial intermediaries similar to that exercised

over the commercial banks. According to him, there is little reason for believing that the central bank's control is weakened by the presence of financial intermediaries. Moreover, so long as the public does not switch easily from bank deposits into indirect securities of intermediaries, the presence of intermediaries may increase the leverage of the central bank on economic activity. It implies that it is not possible for the interest rates to settle back at their old levels even by the operation of non-bank financial intermediaries. Rather, interest rates would tend to rise further. The money supply will remain tight and its influence on spending would be restrictive.

2. Further, the rapid growth of NBFIs has helped to strengthen the effectiveness of monetary policy rather than weaken it where the NBFIs have been controlled. But Gurley and Shaw do not elaborate how they should be controlled.

Conclusion

Despite the weaknesses of both Radcliffe Report and Gurley-Shaw views, they highlight the role of NBFIs in creating liquidity which affects aggregate demand and economic activity. They emphasise that the success of monetary policy depends not in controlling the money supply but general liquidity. Thus the liquidity theory of money provides a new and realistic dimension to monetary policy.

EXERCISES

1. Explain briefly the liquidity theory of money.
2. What do you mean by liquidity? Explain the Radcliffe-Sayers view about the liquidity theory of money.
3. Explain the Gurley-Shaw view about the liquidity theory of money.
4. Explain the liquidity approach to monetary policy.

CHAPTER 49

FISCAL POLICY

MEANING

Fiscal policy means the use of taxation and public expenditure by the government for stabilisation or growth. According to Culberston, “By fiscal policy we refer to government actions affecting its receipts and expenditures which are ordinarily taken as measured by the government's receipts, its surplus or deficit."[1] The government may offset undesirable variations in private consumption and investment by compensatory variations of public expenditures and taxes. Arthur Smithies defines fiscal policy as "a policy under which the government uses its expenditure and revenue programmes to produce desirable effects and avoid undesirable effects on the national income, production and employment." Though the ultimate aim of fiscal policy is the long-run stabilisation of the economy, yet it can be achieved by moderating short-run economic fluctuations. In this context, Otto Eckstein defines fiscal policy as "changes in taxes and expenditures which aim at short-run goals of full employment and price-level stability."[2]

1. J.M. Culberston, *Macro Economic Theory and Stabilisation Policy*, 1968.

2. O. Eckstein, *Public Finance*, (2nd ed.), 1967.

OBJECTIVES OF FISCAL POLICY

The following are the objectives of fiscal policy:

1. To maintain and achieve full employment.

2. To stabilise the price level.

3. To stabilise the growth rate of the economy.

4. To maintain equilibrium in the balance of payments.

5. To promote the economic development of underdeveloped countries.

INSTRUMENTS OF FISCAL POLICY

Fiscal policy through variations in government expenditure and taxation profoundly affects national income, employment, output and prices. An increase in public expenditure during depression adds to the aggregate demand for goods and services and leads to a large increase in income via the multiplier process; while a reduction in taxes has the effect of raising disposable income thereby increasing consumption and investment expenditure of the people. On the other hand, a reduction of public expenditure during inflation reduces aggregate demand, national income, employment, output and prices; while an increase in taxes tends to reduce disposable income and thereby reduces consumption and investment expenditures. Thus the government can control deflationary and inflationary pressures in the economy by a judicious combination of expenditure and taxation programmes. We discuss the following instruments of fiscal policy.

Compensatory Fiscal Policy

The compensatory fiscal policy aims at continuously compensating the economy against chronic tendencies toward inflation and deflation by manipulating public expenditure and taxes. It, therefore, necessitates the adoption of fiscal measures over the long-run rather than once-for-all measures at a point of time. When there are deflationary tendencies in the economy, the government should increase its expenditure through deficit budgeting and reduction in taxes. This is essential to compensate for the lack in private investment and to raise effective demand, employment,

output and income within the economy. On the other hand, when there are inflationary tendencies, the government should reduce its expenditure by having a surplus budget and raising taxes in order to stabilise the economy at the full employment level. The compensatory fiscal policy has two approaches: (1) built-in stabilisers; and (2) discretionary fiscal policy.

(1) Built-in Stabilisers

The technique of built-in flexibility or stabilisers involves the automatic adjustment of expenditure and taxes in relation to cyclical upswings and downswings within the economy without deliberate action on the part of the government. Under this system, changes in the budget are automatic and hence this technique is also known as one of automatic stabilisation. The various automatic stabilisers are corporate profits tax, income tax, excise taxes, old age, survivors and unemployment insurance and unemployment relief payments. As instruments of automatic stabilisation, taxes and expenditure are related to national income. Given an unchanged structure of tax rates, tax yields vary directly with movements in national income, while government expenditure varies inversely with variations in national income. In the downward phase of the business cycle when national income is declining, taxes which are based on a percentage of national income automatically decline, thereby reducing the tax yield. At the same time, government expenditures on unemployment relief and social security benefits automatically increase. Thus there would be an automatic budget deficit which would counteract deflationary tendencies. On the other hand, in the upward phase of the business cycle when national income is rising rapidly, the tax yield would automatically increase with the rise in tax rates. Simultaneously, government expenditures on unemployment relief and social security benefits automatically decline. These two forces would automatically create a budget surplus and thus inflationary tendencies would be controlled automatically.

Its Merits

Built-in stabilisers have certain advantages as a fiscal device. They are:

1. The built-in stabilisers serve as a cushion for private purchasing power

when it falls and lessen the hardships on the people during deflationary period.

2. They prevent national income and consumption spending from falling at a low level.

3. There are automatic budgetary changes in this device and the delay in taking administrative decisions is avoided.

4. Automatic stabilisers minimise the errors of wrong forecasting and timing of fiscal measures.

5. They integrate short-run and long-run fiscal policies.

Its Limitations

It has the following limitations :

1. The effectiveness of built-in stabilisers as an automatic compensatory device depends on the elasticity of tax receipts, the level of taxes and flexibility of public expenditures. The greater the elasticity of tax receipts, the greater will be the effectiveness of automatic stabilisers in controlling inflationary and deflationary tendencies. But the elasticity of tax receipts is not so high as to act as an automatic stabiliser even in advanced countries like America.

2. With low level of taxes even a high elasticity of tax receipts would not be very significant as an automatic stabiliser doing a downswing.

3. The built-in stabilisers do not consider the secondary effects of stabilisers on after-tax business incomes and of consumption spending on business expectations.

4. This device keeps silent about the stabilising influence of local bodies, state governments and of the private sector economy.

5. They cannot eliminate the business cycles. At the most, they can reduce their severity.

6. Their effects during recovery from recession are unfavourable. Economists, therefore, suggest that built-in stabilisers should be supplemented by discretionary fiscal policy.

(2) Discretionary Fiscal Policy

Discretionary fiscal policy requires deliberate change in the budget by such actions as changing tax rates or government expenditure or both. It may generally take three forms: (i) changing taxes with government expenditure constant, (ii) changing government expenditure with taxes constant, and (iii) variations in both expenditure and taxes simultaneously.

(i) When taxes are reduced, while keeping government expenditure unchanged, they increase the disposable income of households and businesses. This increases private spending. But the amount of increase will depend on whom the taxes are cut, to what extent, and on whether the taxpayers regard the cut temporary or permanent. If the beneficiaries of tax cut are in the higher middle income group, the aggregate demand will increase much. If they are businessmen with little incentive to invest, tax reductions are temporary. This policy will again be less effective. So this is more effective in controlling inflation by raising taxes because high rates of taxation will reduce disposable income of individuals and businesses thereby curtailing aggregate demand.

(ii) The second method is more useful in controlling defationary tendencies. When the government increases its expenditure on goods and services, keeping taxes constant, aggregate demand goes up by the full amount of the increase in government spending. On the other hand, reducing government expenditure during inflation is not so effective because of high business expectations in the economy which are not likely to reduce aggregate demand.

(iii) The third method is more effective and superior to the other two methods in controlling inflationary and deflationary tendencies. To control inflation, taxes may be increased and government expenditure be raised to fight depression.

Its Limitations

This policy has the following limitations:

1. The discretionary fiscal policy depends upon proper timing and accurate forecasting. Accurate forecasting is essential to judge the stage of cycle through which the economy is passing. It is only then that appropriate fiscal action can be taken. Wrong forecasting may accentuate rather than moderate the cyclical swings. Economics is not an exact science in correct forecasting. As a result, fiscal action always follows after the turning points in business cycles.

2. There are delays in proper timing of public spending. In fact, discretionary fiscal policy is subject to three time lags.

(i) There is the "decision lag" the time required in studying the problem and taking the decision. The lag involved in this process may be too long.

(ii) Once the decision is taken, there is an "execution lag." It involves expenditure which is to be allocated for the execution of the programme. In a country like the USA, it may take two years and less than a year in the U.K.

(iii) Certain public work projects are so cumbersome that it is not possible to accelerate or slow them down for the purpose of raising or reducing spending on them.

Conclusion

Despite the higher multiplier effect of government spending as against changes in tax rates, the latter can be operated more promptly than the former. Emphasis has thus shifted to taxation as the best fiscal device for controlling cyclical fluctuations. Thus when the turning point of a business cycle is already underway, discretionary fiscal action tends to strengthen the built-in stabilisers, as has been the experience of developed countries like the USA.

Budgetary Policies—Countercyclical Fiscal Policy

The budget is the principal instrument of fiscal policy. Budgetary policies

exercise control over the size and relationship of government receipts and expenditures. We discuss below the common budgetary policies that can be adopted for stabilising the economy.

(1) Budget Deficit—Fiscal Policy during Depression.

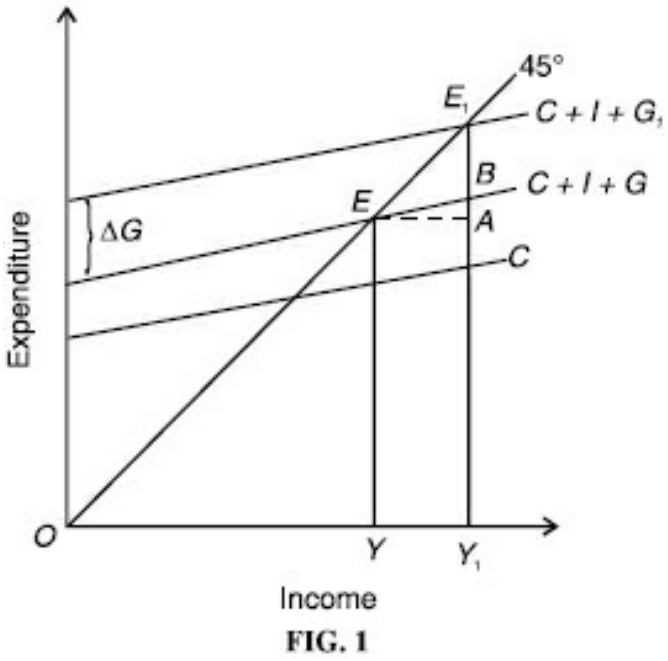


FIG. 1

Deficit budgeting is an important method of overcoming depression. When government expenditures exceed receipts, larger amounts are put into the stream of national income than they are withdrawn. The deficit represents the net expenditure of the government which increases national income by the multiplier times the increase in net expenditure. If the *MPC* is 2/3, the multiplier will be 3; and if the net increase in government expenditure is Rs. 100 crores, it will increase national income to Rs. 300 crores (= 100 × 3). Thus the budget deficit has an expansionary effect on aggregate demand whether the fiscal process leaves marginal propensities unchanged or whether a redistribution of disposable receipts occurs. The expansionary effect of a budget deficit is shown diagrammatically in Figure 1 where *C* is the consumption function. $C + I + G$ represents consumption, investment and government expenditure (the total spendings function) before the budget is introduced. Suppose government expenditure of ΔG is injected into the economy. As a result, the total spendings function shifts upward to $C + I + G_1$. Income increases from *OY* to OY_1 when the equilibrium position moves from *E* to E_1. The increase in income YY_1 (= $EA = E_1A$) is greater than the increase in government expenditure E_1B (= ΔG). *BA* ($E_1A - E_1B$) represents increase in consumption. Thus the budget deficit is always expansionary, the rise in national income being (YY_1) greater than the actual amount of government

spending ($\Delta G = E_1B$). In this method of budget deficit, taxes are kept intact.

Budget deficit may also be secured by reduction in taxes and without government spending. Reduction in taxes tends to leave larger disposable income in the hands of the people and thus stimulates increase in consumption expenditure. This, in turn, would lead to increase in aggregate demand, output, income and employment. This is illustrated in Figure 2, where C is the original consumption function. Suppose tax is reduced by ET, it will shift the consumption function upward to C_1. Income will increase from OY to OY_1.

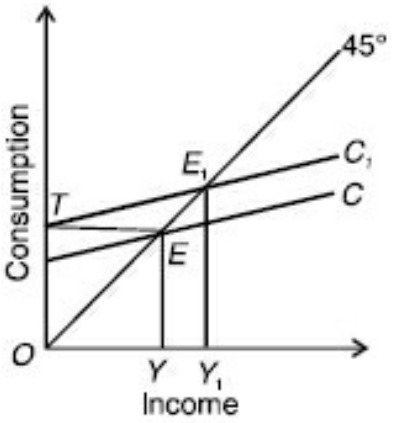


FIG. 2

However, reduction in taxes is not so expansionary via increased consumption expenditure because the tax relief may be saved and not spent on consumption. Businessmen may not also invest more if business expectations are low. Therefore, to safeguard against such eventualities, the government should follow the policy of reduction in taxes with increased government spending. Its multiplier effect will be much higher in case we also assume that some consumption and investment expenditures increase due to tax relief.

(2) Surplus Budget—Fiscal Policy during Boom

Surplus in the budget occurs when the government revenues exceed expenditures. The policy of surplus budget is followed to control inflationary pressures within the economy. It may be through *increase in taxation or reduction in government expenditure* or *both*. This will tend to reduce income and aggregate demand by the multiplier times the reduction in government or/and private consumption expenditure (as a result of increased taxes). This is explained with the aid of Figure 1, where the economy is at the initial equilibrium position E_1. Suppose the government expenditure is reduced by ΔG so that the total spending function $C + I + G_1$ shifts downward to $C + I + G$. Now E is the new

equilibrium position which shows that the income has declined to OY from OY_1 as a result of reduction in government expenditure by E_1B. The fall in income Y_1Y (= AE) > E_1B the reduction in expenditure because consumption has also been reduced by BA.

There may be *budget surplus without government spending when taxes are raised*. Enchanced taxes reduce the disposable income with the people and encourage reduction in consumption expenditure. The result is fall in aggregate demand, output income and employment. This is illustrated in Figure 3, where C is the consumption function before the imposition of the tax. Suppose a tax equal to ET is introduced. The consumption function shifts downward to C_1. The new equilibrium position is E_1. As a result, income falls from OY to OY_1.

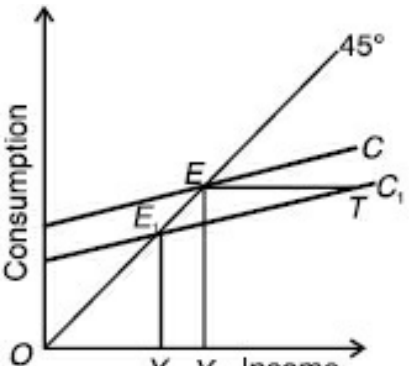


FIG. 3

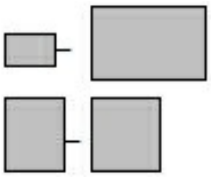
(3) Balanced Budget

Another *expansionist* fiscal policy is the balanced budget. In this policy the increase in taxes (ΔT) and in government expenditure (ΔG) are of an equal amount. This has the impact of increasing *net* national income. This is because the *reduction in consumption resulting from the tax is not equal to the government expenditure.*

The basis for the expansionary effect of this kind of balanced budget is that a tax merely tends to reduce the level of disposable income. Therefore, when only a portion of an economy's disposable income is used for consumption purposes, the economy's *consumption expenditure* will not fall by the full amount of the tax. On the other hand, *government expenditure* increases by the full amount of the tax. Thus the government expenditure rises more than the fall in consumption expenditure due to the tax and there is *net* increase in national income.

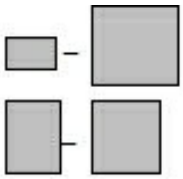
The balanced budget theorem is based on the combined operation of the tax multiplier and the government-expenditure multiplier. In this, the tax

multiplier is smaller than the government-expenditure multiplier. The government-expenditure multiplier is

or  ...(1)

which indicates that the change in income (ΔY) will equal the multiplier (1/1– *c*) times the change in autonomous government expenditure.

The tax multiplier is

or  ...(2)

which shows that the change in income (Δ*Y*) will equal multiplier (1/1– *c*) times the product of the marginal propensity to consume (*c*) and the change in taxes (Δ*T*).

A simultaneous change in public expenditure and taxes may be expressed as a combination of equations (1) and (2). Thus the balanced budget multiplier

 or $k_b - 1$

Since Δ*G* = Δ*T*, income will change by an amount equal to the change in government expenditure and taxes.

To understand it, it is explained numerically. Suppose the value of *c* = 2/3 and the increase in government expenditure Δ*G* = Rs 10 crores. Since Δ*G* = Δ*T*, therefore the increase in taxes (lumpsum) Δ*T* = Rs. 10 crores.

We first calculate the government-expenditure multiplier,

The tax multiplier is $k_T = \frac{\Delta Y}{\Delta T} = \frac{-c}{1-c} = \frac{-2/3}{1-2/3} = -2$

To arrive at the increase in income as a result of the combined operation of the government expenditure multiplier and the tax multiplier, we write the balanced budget multiplier equation as

and fit in the above values of c, ΔG and ΔT so that

$$k_b = \Delta Y = 3\Delta G - 2\,\Delta T$$

$$= 3 \times 10 - 2 \times 10 = \text{Rs. } 10 \text{ crores}$$

Thus the increase in income (ΔY) exactly equals the increase in government expenditure (ΔG) and the lumpsum tax (ΔT) i.e., Rs. 10 crores. Hence k_b=1.

This balanced budget multiplier or *unit multiplier* is explained with the help of Figure 4. C is the consumption function before the imposition of the tax with income at OY_0. level. Tax of AG amount is imposed. As a result, the consumption function shifts downward to C_1. Now government expenditure of GE amount is injected into the economy which is equal to the tax yield AG. The new government expenditure line is $C_1 + G$ which determines OY income at point E. The increase in income Y_0Y equals the tax yield AG and the increase in government expenditure GE. This proves that income has risen by 1 (one) times the amount of increase in government exenditure which is a balanced budget expansion.

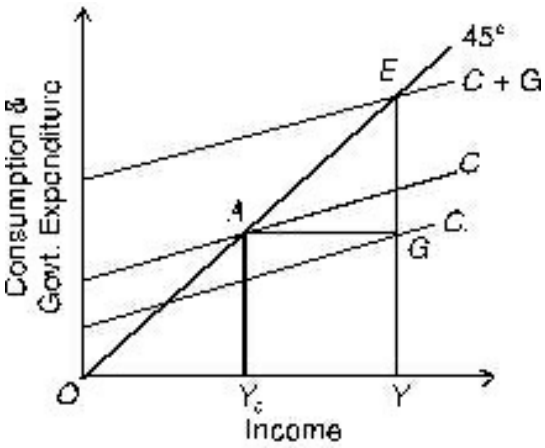


FIG. 4

CROWDING OUT AND FISCAL POLICY

The term *crowding out* refers to the reduction in private expenditure (or investment) caused by an increase in government expenditure through deficit budget via a tax cut or increased money supply or bond issue. An increase in government expenditure raises aggregate demand, national income and interest rates thereby reducing private investment. This is called the *crowding effect* of fiscal policy.

The Keynesians and monetarists differ on the effects of budget deficit on the crowding out effect. The main difference between the two arises from the fact that the Keynesians emphasise on "first-round" (short-run) effects which show "once-for-all shift" of the *IS* curve, whereas the monetarists emphasise the "ultimate" (long-run) effects.

The Keynesian crowding out theory states that when the government resorts to deficit financing by issuing new bonds, its spending increases. National income rises. If the money supply is held constant, people will need more money for business which will raise the rate of interest. A higher rate of interest will crowd out (reduce) private investment spending. These are the first-round effects which are explained in Figure 5, where E_1 is the initial equilibrium position. The rise in government expenditure financed by issuing bonds shifts the IS_1 curve rightward to IS_2 on a "once-for-all" basis and it cuts the *LM* curve at point E_2. Since the money supply is constant, E_2 is the new equilibrium level of the economy. The multiplier process raises the income level from Y_1 to Y_2 and the interest rate from R_1 to R_2. Higher interest rate crowds out a certain amount of private investment.

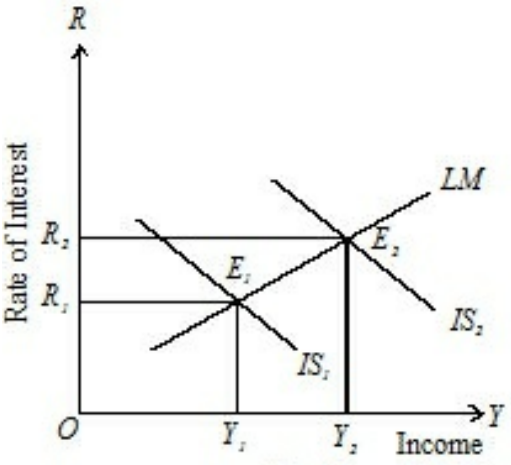


Fig. 5

The Keynesians hold that a deficit financed by printing notes (money creation) is more expansionary than bond-financed. But they do not

believe that the reduction in private expenditure caused by a higher interest will completely offset the increased government expenditure. In other words, the crowding out of private investment will not be full. The reason for this is that a high interest rate has dual effects. *First*, it reduces private spending. *Second*, a high interest rate leads people to economise on cash balances. They, therefore, divert idle cash holdings for transactions purposes. That is why crowding out of private investment is only partial. On the other hand, Friedman exphasises the *ultimate effects* of a budget deficit (whether bond-financed or money financed) by taking account of the *wealth effect*. When the government increases its expenditure by selling bonds in the market, their buyers feel themselves wealthier than before. The reason is that they expect to have more resources available for consumption and other purposes in the future. As a result, they tend to increase the demand for money which shifts the *LM* curve *leftward*. This analysis assumes that bonds issued by the government are considered on wealth. Further, both the demand for money and expenditure on consumption are positively related to wealth.

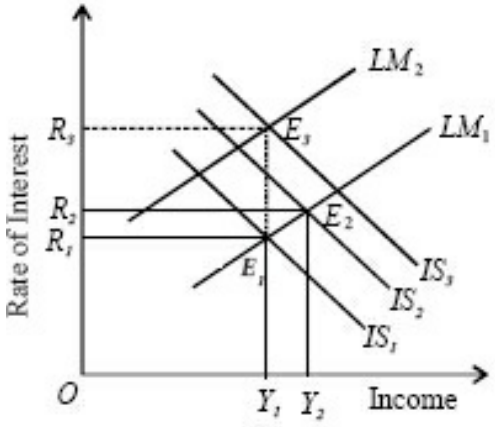


Fig. 6

Suppose the government increases its expenditure with *bond-financed* budget deficit. As a result, the public expenditure on buying bonds also increases.The rise in public expenditure shifts the IS_1 curve rightward to IS_2 in Figure 6 on a "once-for-all basis". This "first-round" effect raises the level of national income from Y_1 to Y_2, given the LM_1 curve. The increase in national income, in turn, raises the demand for money and the purchase of government bonds by the public further raise the demand for money due to the wealth effect. As the LM_1 curve shifts leftward to LM_2 and the IS_2 curve shifts rightward to IS_3 so that the ultimate equilibrium E_3 is established at the initial level of income Y_1. According to Friedman, the rise in interest rate to R_3 reduces private investment so that bond-financed government expenditure crowds out private investment. But the total expenditure remains unchanged and *fiscal policy has no expansionary effect on national income.*

If the budget deficit is *money-financed*, it will have an expansionary effect. This is because the increase in money supply is greater than the wealth effect on the demand for money. In this case, the LM_1 curve shifts rightward to LM_2, as shown in Figure7. The increase in government expenditure shifts the IS_1 curve rightward to IS_2. The first-round effect raises the level of income from Y_1 to Y_2. According to Friedman, in a money-financed deficit, the money stock continues to grow and the *LM* curve continues to shift to the right causing falling interest rates. In this case, the *LM* curve exerts a dominance influence on subsequent changes in income than the *IS* schedule. The ultimate (long-run) equilibrium is shown with the shifting of the IS_2 curve rightward to IS_4 and also of the LM_2 curve rightward to LM_4 so that Y_4 equilibrium income level is established. The rate of interest has fallen from Y_2 E_2 to Y_4 E_4. Thus money-financed deficit is expansionary and it does not crowd out private investment.

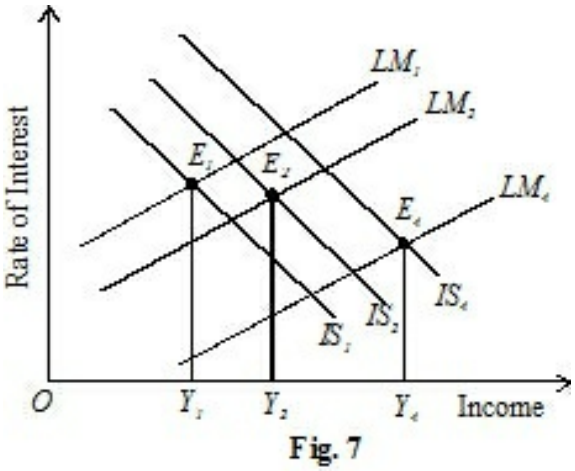


Fig. 7

Blinder and Solow have criticised Friedman's crowding out model of debt-financed deficit for ignoring interest payments on outstanding debt. They point out that the government has not only to finance the budget deficit but also interest payments on outstanding debt. They have shown that if private expenditure and demand for money are subject to wealth effects, then the IS and *LM* curves will be shifting from period to period and the short-run equilibrium will differ from the long-run equilibrium depending upon whether the budget is bond-financed or money-financed.

The short-run and long-run equilibrium situations in the case of *bond-financed* budget deficit are shown in Figure 8 (A). The rise in government expenditure as a result of bond-financed deficit shifts the IS_1 curve *rightward* to IS_2. This shift is due to both the increase in government expenditure and rise in private expenditure following the wealth effect of

bonds. The LM_1 curve shifts leftward to LM_2 as a result of wealth effect which increases the demand for money. This raises the short-run equilibrium level of income from Y_1 to Y_2.

If the budget deficit is *money financed* creation, the increase in government expenditure is once-for-all increase in the short-run so that the IS_1 curve shifts rightward to IS_2 by the same extent, in Panel (B) of the figure. But the increase in the supply of money being greater than the wealth-induced increase in the demand for money, the LM'_1 curve shifts rightward to LM'_2 in Panel (B). This raises the short-run equilibrium level of income from Y'_1 to Y'_2. A comparison of the bond-financed and money-financed situations shows that money-financed income level Y'_2 is greater than the bond-financed level Y_2. This is because the increase in money supply lowers the interest rate form Y'_1 E_1 to Y'_2 E_2, in Panel (B).

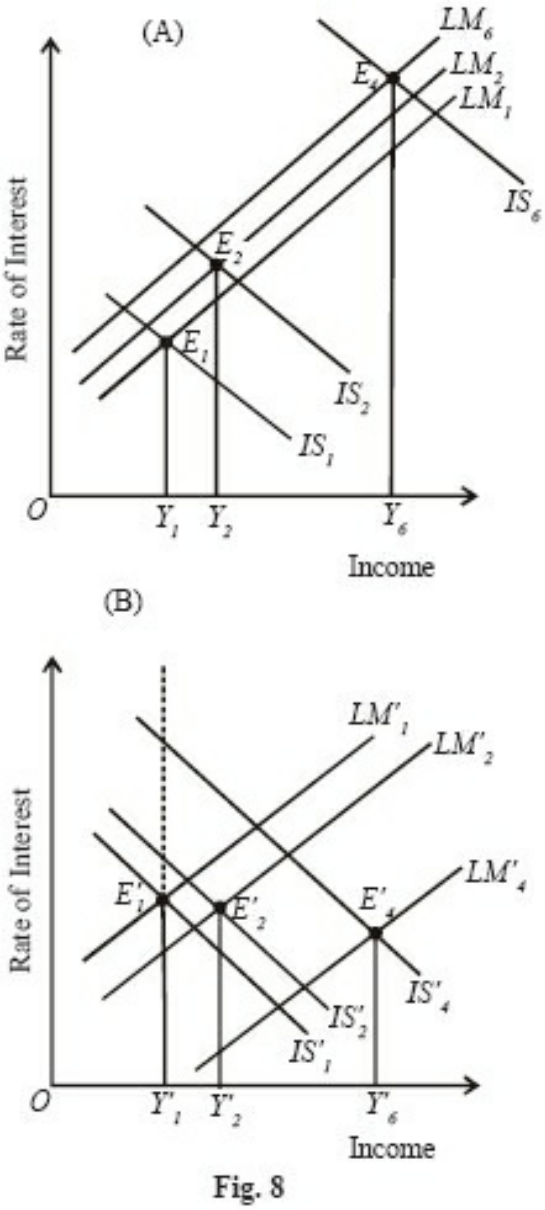


Fig. 8

Hence money-financed deficit is more expansionary and it does not crowd out private investment. On the other hand, the bond-financed deficit raises the interest rate from Y_1 E_1 to Y_2 E_2 when the national income rises to Y_2 in Panel (A). It is also expansionary but it crowds out a part of private investment.

In the long-run, bond-financing is more expansionary than money-financing. This is because when "deficits are bond-financed, income must rise sufficiently to produce tax receipts (at given tax rates) that not only match the increased government expenditure on goods and services, but

also cover the interest payments on the increased government debt. If deficits are, on the other hand, money financed, long-run equilibrium is established when income has merely risen sufficiently to produce tax revenues that each match the increased expenditure on goods and services." Figure 8 (A) shows the bond-financed situation. When in the long-run the IS_2 curve shifts to IS_6 and the LM_2 curve to LM_6, the new equilibrium level of income is set at Y_6. The money-financed situation is shown in Panel (B) where in the long-run the IS'_2 curve shifts to IS'_4 and LM'_2 curve to LM'_4 and the new equilibrium is established at Y'_4 income level. As is clear from the two figures, Y_6 income level is greater than Y'_4 level.

The above analysis shows that the long-run effect of increased bond-financed government deficit is more expansionary but it crowds out private investment because the rate of interest rises sufficiently high. But the money-financed deficit though less expansionary than the former, it does not crowd out private investment through wealth effects.

EXERCISES

1. What do you mean by fiscal policy ? Explain the fiscal measures which are undertaken by the government.
2. Discuss compensatory fiscal policy.
3. Discuss the budgetary policies that should be adopted during depression and boom.
4. Write notes on (i) balanced budget (ii) built-in stabilisers; (iii) discretionary fiscal policy.
5. What is crowding out ? Explain the crowding out effects of fiscal policy.

CHAPTER 50

MONETARISM VERSUS KEYNESIANISM

INTRODUCTION

Monetarism refers to the followers of Milton Friedman who hold that "*only money matters*", and as such monetary policy is a more potent instrument than fiscal policy in economic stabilisation. On the other hand, Keynesianism refers to the followers of Keynes who believe that "*money does not matter*," and for economic stabilisation fiscal policy is a more powerful tool than monetary policy. The adherents of monetarism are known as the monetarists and of Keynesianism as the fiscalists. We discuss below the views of the monetarists and fiscalists about the causes of changes in national income and the roles of monetary and fical policies in economic stabilisation.

THEORETICAL DIFFERENCES

The monetarists emphasise the role of money in explaining short-term changes in national income. They argue that the role of money has been neglected by Keynesians, if not by Keynes himself. Friedman and Schwartz[1] have shown that changes in the money supply cause changes in national income. The monetarists believe that all recessions and depressions are caused by severe contraction of money and credit, and booms and inflations by excessive increases in the money supply. The Keynesians reject the monetarists view that changes in national income are caused solely by changes in the money supply. Rather they hold that changes in national income cause changes in the supply of money. The moderate Keynesians still believe like monetarists that hyper-inflations

are caused by excessive money supply. On the other hand, the extreme Keynesians hold that non-monetary factors like investment cause depressions and booms.

The Monetarist View

According to the monetarists, the money supply is the "dominant, though not exclusive" determinant of both the level of output and prices in the short run, and of the level of prices in the long run. The long-run level of output is not influenced by the money supply. It is dependent on such real factors as technology and the quantity and quality of productive resources. Although factors other than the money supply affect the level of output, employment and prices in the short run, yet their effects are subordinate to that of the money supply. A change in the money supply will inevitably affect the price level and output in the short run. But in the long run, the effect of change in the money supply will be entirely on the price level. Because the economy is near full employment level in the long run, the increase in national income will consist mainly of higher prices. Thus changes in the money supply affect national income directly. This is due to the assumption that the velocity of circulation of money is stable.

The Keynesians, on the other hand, maintain that changes in the money supply work *indirectly* on the level of aggregate expenditure and income through changes in interest rates. The main difference between the monetarists and Keynesians rests on the influence of interest rates on the demand for and supply of money and aggregate expenditure.

The monetarists hold that the rate of interest plays no part in determining the demand for money. The demand for money is the transactions demand for money which is determined by the level of income. The monetarists believe that the money supply is also not influenced by interest rates. This insensitiveness of both demand for and supply of money is based on the quantity theory of money. The simplest quantity equation is $MV = PQ$, where M is quantity of money, V is its velocity, Q is the number of physical transactions, and P is the price level. Taking $PQ = Y$, where Y is the national income, the quantity equation becomes $MV = Y$. As V remains constant, changes in M cause changes in Y.

1. M. Friedman and Anna Schwartz, *A Monetary History of the United States,* 1957-1960, 1963.

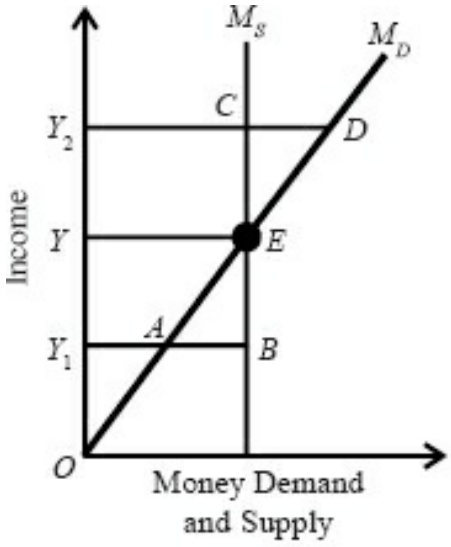


Fig. 1

Figure 1 the explains the monetarist view graphically. The demand for and supply of money are taken on the horizontal axis and income on the vertical axis. The demand curve M_D is drawn from the origin which shows the amount of money people want to hold for transactions purposes at various levels of income. M_S is the money supply curve which is drawn vertical to show a given supply of money. *OY* is the equilibrium level of income where the money demanded and supplied equal at *E*. If people have more money than they want equal to *AB* at OY_1 income level, they will start spending it till income increases to the equilibrium level *OY*. If they have less money than they want equal to *CD* at OY_2 income level, they will start reducing their spending till it reaches the equilibrium level *OY*.

The change in the money supply affects aggregate demand and income through effects on a wide range of assets than "the bonds only" model of the Keynesians. This view of the monetarists is based on the belief that money is a good substitute for all types of assets such as securities, houses, durable consumer goods, etc. Friedman's historical findings show a "*stable money demand function*" which implies that the demand for money is a stable function of peoples' income. In other words, the amount of money people want to hold is related in a fixed way to their income. If the central bank increases the money supply, it affects interest rates in three different ways. *First*, there is the *liquidity effect* which causes a very short-run reduction in interest rates. *Second*, as a result of liquidity effect, people will sell securities and their holdings of money will increase. They will, therefore, spend their excess money balances on financial assets and on durable consumer goods, houses, etc. This increase in aggregate expenditure on assets and goods will tend to raise output, employment and income. This is the *output effect*. This will lead to a rise in interest rates because of the rise in output and demand for money resulting from the

liquidity effect. *Finally*, there is the *price expectations effect* which occurs due to the expectations of lenders that inflation will continue. They will demand higher interest rates in order to cover the expected inflation rate. Thus with the short-run liquidity effect bringing a downward pressure on interest rates and both the output and price expectations effects bringing an upward pressure on them, the combined effect will be an increase in interest rates. These will, in turn, discourage investment, and reduce output and employment. This is illustrated in Figure 2.

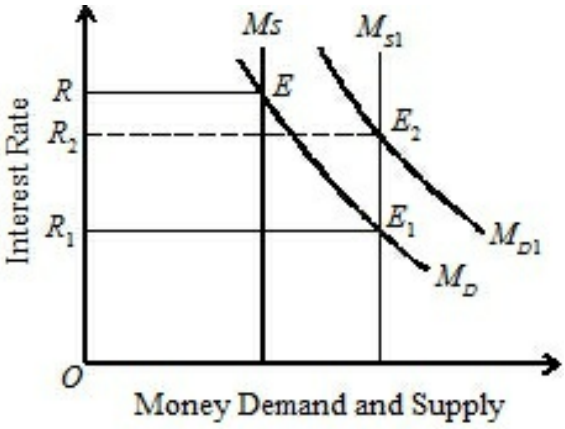


Fig. 2

The demand for and supply of money are taken on the horizontal axis and the interest rate on the vertical axis. M_D is the demand curve for money and M_S the money supply curve. Both intersect at *E* and determine *OR* interest rate. With the increase in the money supply, the M_S curve shifts to M_{S1}. If the money demand remains constant, the interest rate would fall from *OR* to OR_1. The increased demand for financial assets and durable consumer goods leads to an increase in output and prices so that the demand for money increases at every interest rate. As a result, the M_D curve shifts upward to M_{D1} and intersects the Ms_1 curve at E_2. As a result, the interest rate increases to OR_2. The M_D curve is drawn steeply which reflects the monetarist view that the demand for money is interest elastic.

The Keynesian View

The Keynesians hold just the opposite views to monetarists about the demand for and supply of money and the aggregate expenditure. Both the demand for and supply of money are highly *interest elastic* while the aggregate expenditure is not.

The Keynesians consider the supply of money to be *fixed* in the short run by monetary authorities. The demand for money, also called the liquidity preference, is the desire to hold cash. There are three motives on the part

of the people to hold cash: transactions, precautionary and speculative demand for money. Money held for transactions and precautionary motives is a function of the level of income, and for the speculative motive, it is a function of the interest rate.

To the Keynesians, it is expectations about changes in bond prices or in the market rate of interest that determine the speculative demand for money. The speculative demand for money is a decreasing function of the rate of interest. The higher the rate of interest, the lower the speculative demand for money, and vice-versa. But at a very low interest rate, the speculative demand for money becomes *perfectly elastic*. This is the *liquidity trap* portion of the demand for money curve. In this range, people prefer to keep money is cash rather than invest in bonds because purchasing bonds will lead to loss.

On the whole, given the level of national income, the demand for money is a decreasing function of the interest rate. The Keynesians believe money and financial assets to be good substitutes for each other. They are highly liquid and yield interest. So even small changes in interest rates lead to substitution between money and financial assets. A small fall in the interest rate will mean a rise in the price of securities, which will induce people to sell securities and hold more money. The reverse will be the case in the event of a small rise in the rate of interest. Thus the demand for money is highly interest-elastic under Keynesianism.

The Keynesians believe in the existence of *unemployment equilibrium*. This implies that an increase in money supply can bring about permanent increases in the level of output. The ultimate influence of money supply on the price level depends upon its influence on aggregate demand and the elasticity of the supply of aggregate output.

The Keynesian chain of causation between changes in the quantity of money and in prices is an *indirect* one through the rate of interest. So when the quantity of money is increased, its first impact is on the rate of interest which tends to fall. Given the marginal efficiency of capital, a fall in the rate of interest will increase the volume of investment. The increased investment will raise effective demand through the multiplier effect thereby increasing income, output and employment. Since the

supply curve of factors of production is perfectly elastic in a situation of unemployment, wage and non-wage factors are available at constant rate of remuneration. There being constant returns to scale, prices do not rise with the inrease in output so long as there is any unemployment. Under the circumstances, output and employment will increase in the same proportion as effective demand, and the effective demand will increase in the same proportion as the quantity of money. But "once full employment is reached, output ceases to respond *at all* to changes in the supply of money and so in effective demand. The elasticity of supply of output in response to changes in the supply, which was infinite as long as there was unemployment, falls to zero. The entire effect of changes in the supply of money is exerted on prices, which rise in exact proportion with the increase in effective demand". Thus *so long as there is unemployment, output will change in the same proportion as the quantity of money,* and *there will be no change in prices;* and *when there is full employment, prices will change in the same proportion as the quantity of money*.

In the Keynesian transmission mechanism, changes in the money supply affect aggregate expenditure and national income *indirectly* by changes in interest rates. Suppose the money supply is increased, it lowers the interest rate which, in turn, increases investment and expenditure thereby raising the national income. The mechanism by which changes in the money supply are transmitted into the level of income is the *asset effect*. For example, an increase in the money supply causes people to spend their excess holdings of money on financial assets. This means an increase in the demand for such assets and a rise in their prices. Rise in the prices of assets (securities), brings down the interest rates which, in turn, increase aggregate expenditure, investment and hence income. Thus, according to the Keynesian view, a change in the money supply can only affect aggregate spending and national income first through changes in interest rates, and then only if the aggregate spending is sensitive to interest rate changes.

POLICY DIFFERENCES

Another point of differences between the monetarists and the Keynesians is over the policy prescriptions. According to the monetarists, *monetary*

policy has a greater influence on economic activity than fiscal policy, and fiscal policy is important only in making changes in the money supply. On the other hand, the *Keynesians emphasise the importance of both fiscal and monetary policy* in influencing the economy but they attach more importance to the former than to the latter. First, we study the monetarist view on monetary and fiscal policy and then the Keynesian view.

The Monetarist View

As already analysed above, the monetarists hold that changes in the money supply have a direct influence on aggregate expenditure and thus on income. Let us analyse an expansionary monetary policy followed by monetarists. To begin, suppose the central bank purchased securities in the open market. It raises the price of securities and lowers the rate of interest.* People will, therefore, start selling securities and hold more money. People spend their excess money balances on financial assets and durable consumer goods. Others attracted by low interest rates borrow from banks for expenditure on houses, durable consumer goods, plants and equipments, etc. These forces tend to increase aggregate expenditure and income. This is illustrated in Figure 3. M_D is the demand for money curve and M_S the money supply curve. Both the curves are *interest inelastic*. Initially both are in equilibrium at *OR* interest rate. The increase in money supply to M_{S1} reduces the interest rate from *OR* to OR_1. The demand for money being relatively insensitive to the interest rate, it increases from *OQ* to OQ_1. But the aggregate expenditure is very sensitive to this fall in interest rate so that it increases from *OE* to OE_1 in Panel (B) of the figure where *EC* represents the expenditure curve. Thus an expansionary monetary policy is highly successful in increasing aggregate expenditure and income.

---

* Securities (bonds) prices and interest rates are inversely related.

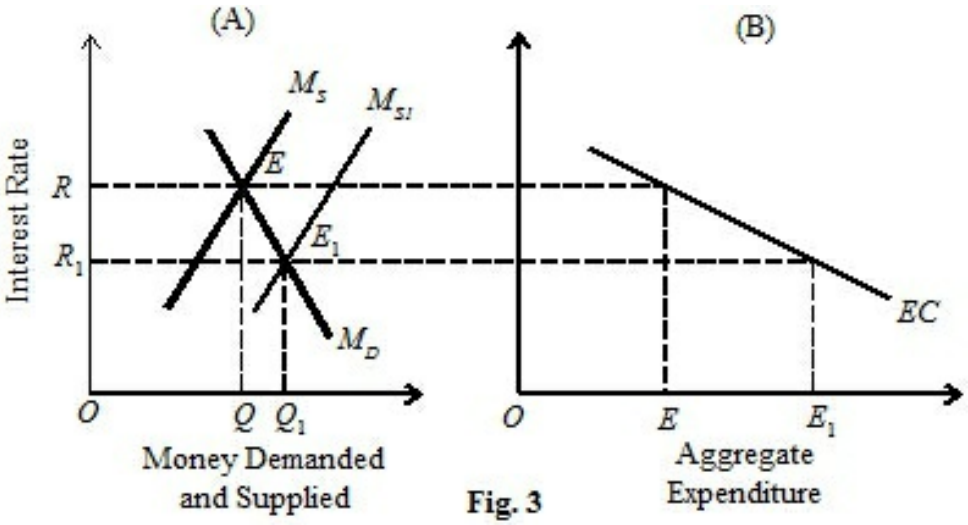


Fig. 3

Now consider *fiscal policy* in Figure 4 (A) and (B). An increase in aggregate expenditure leads to expansion in the economy which shifts the expenditure curve upward from EC to EC_1 in Panel (B) of the figure. As aggregate expenditure incureases from OE to OE_1 national income also rises. This leads to greater demand for money for transactions purposes. In order to meet this increased demand for money, households and firms sell securities which they have and also borrow from banks and other financial institutions. These tendencies have the effect of raising the interest rate. As the demand for money is relatively insensitive to interest rate changes, a very high rate of interest is required to equate the supply of money to the demand for money.

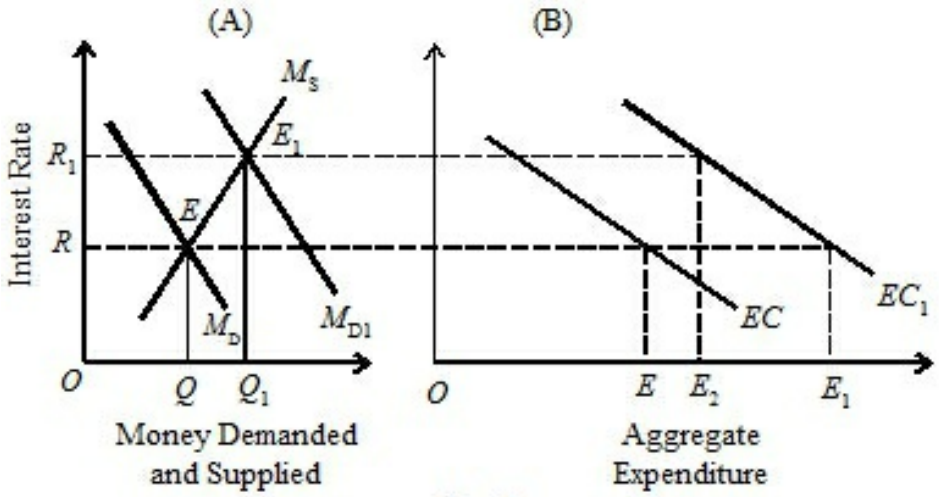


Fig. 4

This is shown in Panel (A) of Figure 4 where the increased demand for money is shown by the M_{D1} curve and the rise in the interest rate by RR_1. But such a large increase in interest rate has the effect of reducing private expenditure from OE_1 to OE_2 so the net effect on aggregate expenditure of an expansionary fiscal policy is only EE_2. Thus according to the monetarists, an increase in government expenditure as a result of an expansionary fiscal policy leads to an increase in government expenditure and income which leads to a large rise in the rate of interest that, in turn, reduces private expenditure almost equal to the increase in government expenditure.

The Keynesian View

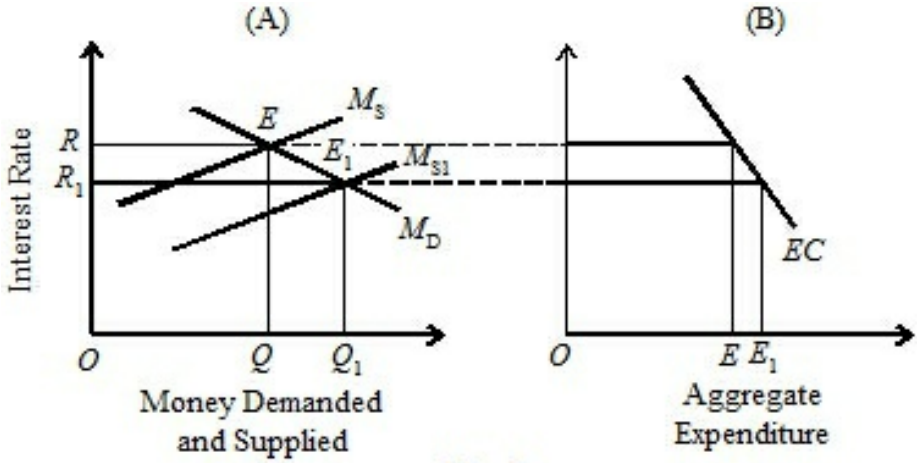


Fig. 5

In contrast to the monetarists, the Keynesians regard monetary policy *relatively less effective* because of the *relative interest inelasticity* of aggregate expenditure. To illustrate, consider an expansionary monetary policy. Suppose the central bank purchases securities in the open market. As a result, the price of securities rises and the interest rate falls. People will, therefore, start selling securities in order to hold more money. As the demand for money is highly interest elastic in the Keynesian system, even a small fall in the rate of interest will induce people to sell securities and hold more money. This is illustrated in Figure 5. Panel (A) of the figure shows that in the initial situation, *OR* is the interest rate and *OQ* the money demanded and supplied. Now the money supply increases due to the central bank's action. This is shown by the Ms_1 curve. When the

money supply rises to OQ_1, the interest rate falls from OR to OR_1. This small fall in the interest rate by RR_1 increases aggregate expenditure by only EE_1, as shown in Panel (B) of the figure. Thus an expansionary policy is not successful in raising aggregate expenditure and income much.

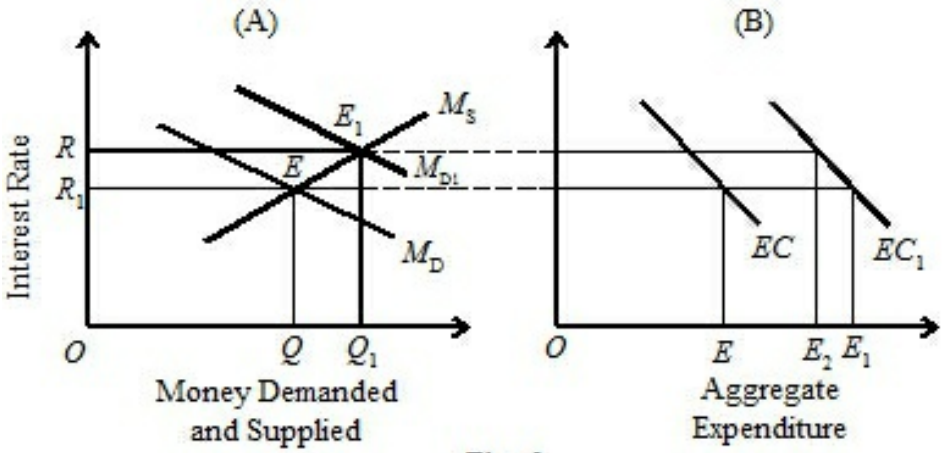


Fig. 6

Though monetary policy is relatively less effective under the Keynesian system, *fiscal policy* is relatively *more effective*. Consider Figure 6. An increase in aggregate expenditure for raising national income leads to expansion in the economy. This shifts the expenditure curve EC upward to EC_1 in Panel (B) of the figure. As aggregate expenditure increases from OE to OE_1, national income rises. This leads to greater demand for money by the people for transactions purposes. To meet this increased demand for money, householders and firms sell securities they possess, and also borrow funds from the banks and other financial institutions. These moves tend to raise the rate of interest. Higher interest rates and the demand for more loans induce commercial banks to reduce their excess reserves, thereby increasing the supply of money from OQ to OQ_1 in Panel (A) of the figure. Further, high interest rates induce bond holders to reduce the amount of money held for speculative purposes because high interest rates mean fall in the price of bonds. As the demand for money is highly interest elastic only a small rise in the interest rate from OR to OR_1 is required to equate the money supply curve M_S with the new money demand curve M_{D1} at E_1 in Panel (A) of Figure 6. But this small increase

in interest rate by RR_1 has a very small effect in reducing private expenditure from OE_1 to OE_2, as shown in Panel (B) of the figure. This is because aggregate expenditure is relatively interest inelastic in the Keynesian system. Thus the effect of an expansionary fiscal policy is net increase in aggregate expenditure by EE_2 and only a small amount of private expenditure is reduced by E_1E_2 due to a rise in the rate of interest by RR_1. Thus the Keynesians regard fiscal policy more effective than monetary policy.

CONCLUSION

The above analysis about monetarism and Keynesianism reveals that both hold almost the opposite views. The monetarists argue that only money matters, and that economic recessions and expansions are caused by decreases and increases of the money supply. They, therefore, advocate control of the money supply to stabilise cyclical fluctuations. They emphasise that the growth rate of money is the principal determinant of the behaviour of national income. This view is based on a number of historical studies carried out by Friedman and Schwartz, Friedman and Meiselman,[2] and Anderson and Jordan[3]. These studies reveal that there is a very close relation between money supply and national income than between national income and any of the Keynesian variables like aggregate expenditure.

Though the monetarists have tried to enforce their position on the basis of empirical studies yet they are themselves skeptical about the success of monetary policy in contrast to fiscal policy. They agree that as economic stabiliser, monetary policy may do more harm than good because of the operation lag. The operation lag refers to the time elapsing between the taking of action and the effective impact of that action on the economic situation. On the average, it takes a long time for a change in the money supply to affect national income, so the operation lag is long. Friedman himself admits that the time lag involved is so large that contracyclical monetary policy might actually have a destabilising effect on the economy. The monetarists, therefore, hold that the economy is basically stable and when disturbed by some change in basic conditions, will

quickly revert to its long-run growth path. That is why, they advocate an annual fixed percentage growth in the money supply and an end to discretion in monetary policy. Friedman, therefore, believes that fiscal policy does not have any potent influence on the economy except that it affects the behaviour of money.

---

2. M. Friedman and David Meiselman, "The Relative Stability of Monetary Velocity and the Investment Multiplier in the United States, 1897-1958," *Stabilization Policies*, 1963.

3. L.C. Anderson and J.L. Jordan, "Monetary and Fiscal Actions: A Test of their Relative Importance in Economic Stabilisation," *Federal Reserve Bank of St. Louis Review*, November 1968.

On the other hand, the Keynesians are not diehards like the monetarists. They take a more reslistic view of monetary and fiscal policy in contrast to the latter. They do not regard the two as competitive but complementary to each other. They do not deny that money does matter, for they believe that monetary policy does influence national income but via changes in the interest rate. But they find monetary policy ineffective in controlling severe depressions and therefore depend upon fiscal policy for this. On the other hand, they combine monetary policy with fiscal policy for controlling booms. In this connection, the views of Walter Heller merit consideration : "*The Keynesian economics*... assigns an important role to both fiscal and monetary policy. Indeed, the appropriate mix of policies has been the cornerstone of the argument...To anyone who fears that the 'new economics' is all fiscal policy, the record offers evidence, and the new economics assurance, that money does matter."[4]

EXERCISES

1. Discuss the views of monetarists and Keynesians on monetary policy.
2. Examine Monetarist and Keynesian controversy about monetary and fiscal policy.
3. Explain the main points of differences between the monetarists and the Keynesian about monetary and fiscal policy.

---

4. Walter W. Heller, *New Dimensions of Political Economy*, 1966 Italics mine.

CHAPTER 51

IS AND LM FUNCTIONS: GENERAL EUILIBRIUM OF PRODUCT AND MONEY MARKETS

INTRODUCTION

This chapter integrates money, interest and income into a general equilibrium model of product and money markets in the Hicks-Hansen [1] diagrammatic framework, known as the *IS-LM* model. The term *IS* is the shorthand expression of the equality of investment and saving which represents the product market equilibrium. On the other hand, the term *LM* is the shorthand expression of the equality of money demand (*L*) and money supply (*M*) and represents the money market equilibrium.

In order to analyse the general equilibrium of product and money markets, it is instructive to study the derivation of the IS and LM functions and their slopes for the understanding of the effectiveness of monetary and fiscal policies.

1. J.R. Hicks, "Mr Keynes and the Classics: A Suggested Interpretation," *Econometrica*, April 1937, reprinted in M.G. Mueller (ed), *Readings in Macroeconomic*, pp. 137-147, and A.H. Hansen, *Monetary Theory and Fiscal Policy*, Ch. 5.

THE PRODUCT MARKET EQUILIBRIUM

The product market is in equilibrium when desired saving and investment are equal. Saving is a direct function of the level of income,

$$S = f(Y) \qquad \ldots(1)$$

Investment is a decreasing function of the interest rate,

$$I = f(r) \qquad \ldots(2)$$

From (1) and (2), we have $S=I$.

The *IS* schedule reflects the equilibrium of the product market. It shows the combinations of interest rate and income levels where saving-investment equality takes place so that the product market of the economy is in equilibrium. It is also known as the "*real sector*" equilibrium.

Deriving the IS curve

The derivation of the *IS* curve is shown in Figure 1. In Panel (A) of this figure, the saving curve *S* in relation to income is drawn in a fixed position on the Keynesian assumption that the rate of interest has little effect on saving. The saving curve shows that saving increases as income increases, viz., saving is an *increasing* function of income. Investment, on the other hand, depends on the rate of interest and the level of income. Given a level of interest rate, the level of investment rises with the level of income. At a 5 per cent rate of interest, the investment curve is I_2 . If the rate of interest is reduced to 4 per cent, the investment curve will shift upward to I_3 . The rate of investment will have to be raised to reduce the marginal efficiency of capital to equality with the lower rate of interest. Thus the

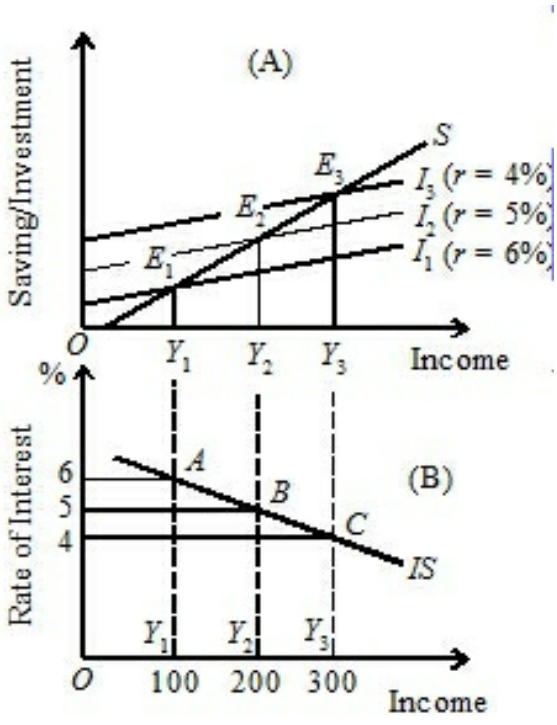


Fig. 1

investment curve I_3 shows more invesment at every level of income. Similarly when the rate of interest is raised to 6 per cent, the investment curve will shift downward to I_1 . The reduction in the rate of investment is essential to raise the marginal efficiency of capital to equal the higher interest rate. In Panel (B) we drive the *IS* curve by marking the level of income at various interest rates. Each point on this *IS* curve represents a level of income at which saving equals investment at various interest rates. The rate of interest is represented on the vertical axis and the level of income on the horizontal axis. If the rate of interest is 6 per cent, the *S* curve intersects the I_1 curve at E_1 in Panel (A) which determines OY_1 income. From this income level which equals Rs 100 crores we draw a dashed line downward to intersect the extended line from 6 per cent at point *A* in Panel (B). At interest rate 5 per cent, the S curve intersects the I_2 curve at E_2 so as to determine OY_2 income (Rs 200 crores). In the lower figure, the point *B* corresponds to 5 per cent interest rate and Rs 200 crores income level. Similarly, the point *C* corresponds to the equilibrium of *S* and I_3 at 4 per cent interest rate. By connecting these points *A*, *B* and *C* with a line, we get the *IS* curve. The *IS* curve in Figure 1(B) slopes *downward* from left to right because as the interest rate falls, investment increases and so do income and saving. In other words, there is a negative relationship between income and interest rate in the real sector of the economy.

The Slope of the IS Curve

This negative slope of the *IS* curve reflects the increase in investment and income as the rate of interest falls. The slope of the *IS* curve depends on two factors: *first* , the sensitiveness (elasticity) of investment and saving to changes in the interest rate, and *second*, on the size of the multiplier.If investment is very sensitive to the rate of interest, the *IS* curve is very *flat*. This is shown by the segment *AB* of the *IS* curve in Figure 2 where a small fall in the rate of interest from R_1 to R_2 leads to a large increase in investment and consequently in saving via proportionately large rise in income from Y_1 to Y_2. The *IS* curve is *interest elastic* in the *AB* segment of the *IS* curve.

On the other hand, if investment is not very sensitive to the rate of interest, the *IS* curve is relatively *steep*. In terms of Figure 2, when the rate of interest falls more from R_2 to R_3, the increase in investment is small and so do saving and income increase by a relatively smaller amount Y_2Y_3. The *BC* segment of the *IS* curve is *less interest elastic*. Any further fall in the rate of interest from R_3 will lead to no change in income because the *IS* curve is vertical in that range. It is *interest inelastic*.

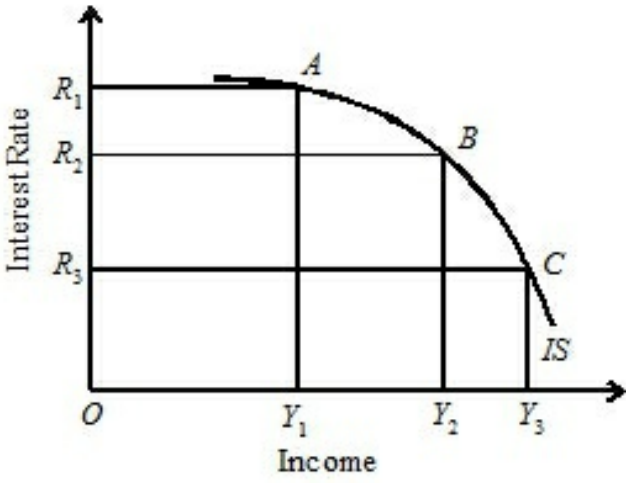


Fig. 2

The shape of the *IS* curve also depends upon the size of the multiplier. If the size of the multiplier is large, the larger is the effect on income of a rise in investment and fall in saving. Thus income is more sensitive to changes in the interest rate and the *IS* curve is flatter.

Shifts in the IS curve

The *IS* function shifts to the right with a reduction in saving. Reduction in saving may be the result of one or more factors leading to increase in consumption. Consumers may like to buy a new product even by reducing saving. The community's wealth may increase due to government's policy and the wealth holders do not like to save the same amount than before. Consumers may start buying more in anticipation of shortages or price rise thereby reducing saving.

The *IS* function also shifts to the right by an autonomous increase in investment. The increase in investment may result from expectations of higher profits in the future, or from innovation, or from expectations concerning increase in the future demand for the product, or from a rise of optimism in general. Moreover, government's expenditure and tax policies have the effect of shifting the *IS* function.

In all these cases, the *IS* function will shift to the right, equal to the decrease in the supply of saving *times* the multiplier or the increase in the investment *times* the multiplier. With the increase in the autonomous investment (or reduction in saving), the *IS* curve shifts from IS_1 to IS_2 and the new equilibrium is established at point E_2 which indicates a higher level of income Y_2 at a higher interest rate R_2, as shown in Figure 3.

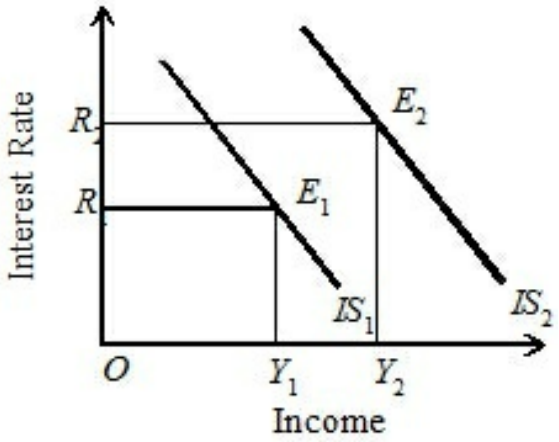


Fig. 3

In the opposite case when investment falls or saving increases, the *IS* function will shift to the left and the equilibrium will be established at a lower level of income and interest rate. This situation can be explained by assuming IS_2, as the original curve.

THE MONEY MARKET EQUILIBRIUM

The money market is in equilibrium when the demand and supply of money are equal. Denoting *L* for money demand and *M* for money supply, the money market is in equilibrium when *L=M*. The demand for money $L=L_T+L_S$ where L_T is the transactions demand for money which is a direct function of the level of income, $L_T=f(Y)$. *Ls* is the speculative demand for money which is a decreasing function of the rate of interest, $L_S=f(r)$. Thus in money market equilibrium, $M=L_T\,(Y)+L_S\,(r)$.

Deriving the LM Curve

The *LM* curve shows all combinations of interest rate and levels of income at which the demand for and supply of money are equal. In other words, the *LM* schedule shows the combinations of interest rates and levels of income where the demand for money (*L*) and the supply of money (*M*) are equal such that the money market is in equilibrium.

The *LM* curve is derived from the Keynesian formulation of liquidity preference schedules and the schedule of supply of money. A family of liquidity preference curves L_1Y_1, L_2Y_2 and L_3Y_3 is drawn at income levels of Rs 100 crores, Rs 200 crores and Rs 300 crores respectively in Figure 4 (A). These curves together with the perfectly inelastic money supply curve *MQ* give us the *LM* curve. The *LM* curve consists of a series of points, each point representing an interest-income level at which the demand for money (*L*) equals the supply of money (*M*). If the income level is Y_1 (Rs 100 crores), the demand for money (L_1Y_1) equals the money supply (*QM*) at interest rate R_1. At the Y_2 (Rs 200 crores) income level, the L_2Y_2 and the *QM* curves equal at R_2 interest rate. Similarly at the Y_3 (Rs 300 crores) income level, the L_3Y_3 and *QM* curves at R_3 interest rate. The supply of money, the liquidity preference, the level of income and the rate of interest provide data for the *LM* curve shown in Figure 4(B).

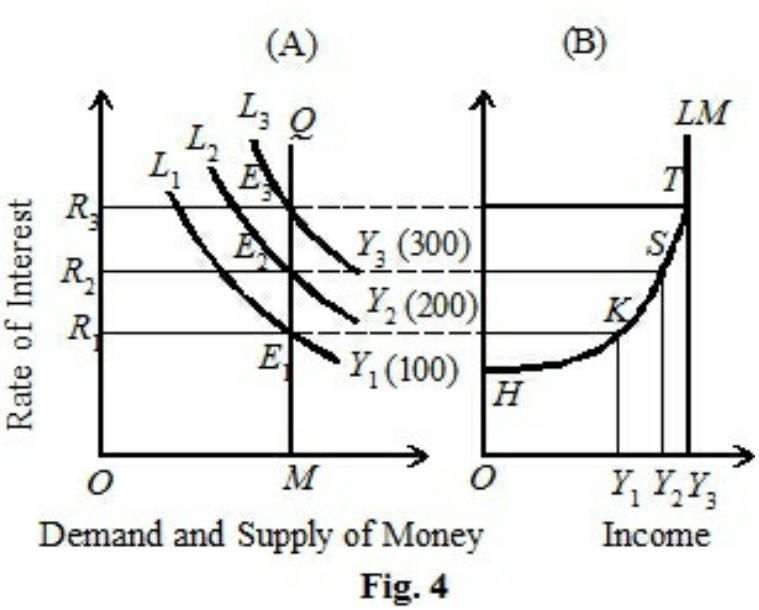


Fig. 4

Suppose the level of income is Y_1 as marked out on the income axis in Figure 4(B). The income of Rs 100 crores generates a demand for money represented by the liquidity preference curve L_1Y_1. From the point E_1 where the L_1Y_1 curve intersects the *MQ* curve, extend a dashed line horizontally to the right so as to meet the line drawn upward from Y_1 and *K* in Figure 4(B). Points *S* and *T* can also be determined in similar manner. By connecting these points *K*, *S* and *T*, we get the *LM* curve. This curve relates different income levels to various interest rates.

The Slope of the LM Curve

The *LM* curve slopes upward from left to right because given the supply

of money, an increase in the level of income increases the demand for money which leads to higher rate of interest. This, in turn, reduces the demand for money and thus keeps the demand for money equal to the supply of money. The smaller the responsiveness of the demand for money to income, and the *larger the responsiveness* of the demand for money to the rate of interest, the *flatter* will be the *LM* curve. This means that a given change in income has a smaller effect on the interest rate.

The *LM* curve is steeper, if a given change in income has a larger effect on the rate of interest. In this situation, the responsiveness of the demand for money to income is larger and is smaller for the interest rate. If the demand for money is *insensitive* to the interest rate, the *LM* curve is *vertical* that is, it is *perfectly inelastic*. This is shown in Panel (B) of Figure 4 as the portion from *T* above on the *LM* curve. In this case, a large change in the interest rate is accompanied by almost no change in the level of income to maintain money market equilibrium. If the demand for money is *very sensitive* to the rate of interest, the *LM* curve is *horizontal*. This is shown by the portion of *LM* curve which starts from *H* on the vertical axis in Panel (B) of Figure 4. The *LM* curve is *perfectly elastic* in relation to the rate of interest. In other words, a small change in the interest rate is accompanied by a large change in the level of income to maintain the money market equilibrium. This portion of the *LM* curve at the extreme left is equivalent to the Keynesian liquidity trap, already explained in the Keynes's theory of interest.

Shifts in the LM Curve

The *LM* function shifts to the right with the increase in the money supply, given the demand for money, or due to the decrease in the demand for money, given the supply of money. If the central bank follows an expansionary monetary policy, it will buy securities in the open market. As a result, the money supply with the public increases for both transactions and speculative purposes. This shifts the *LM* curve to the right.

A decrease in the demand for money means a reduction in the quantity of balances demanded at each level of income and interest rate. Such a decrease in the demand for money balances creates an excess of the

money supplied over the money demanded. This is equivalent to an increase in money supply in the economy which has the effect of shifting the *LM* curve to the right.

This is depicted in Figure 5. With the increase in the money supply, the LM_1 curve shifts to the right as LM_2 which moves the economy to a new equilibrium point E_2. The increase in the money supply brings down the interest rate to R_2 in the money market. This, in turn, increases investment thereby raising the level of income to Y_2.

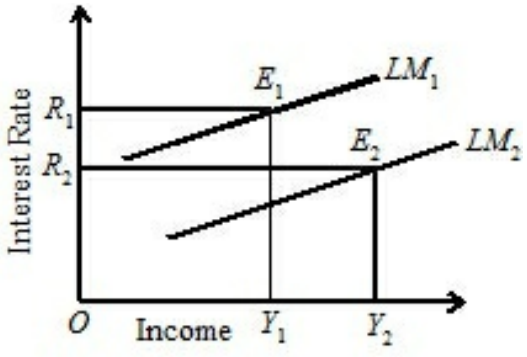


Fig. 5

Contrariwise, a decrease in the money supply, or an increase in the demand for money will shift the *LM* function to the left such that a new equilibrium is established at a higher interest rate and lower income level. This case can be explained by assuming LM_2 as the original curve.

General Equilibrium of Product and Money Market

So far we have analysed the conditions that have to be satisfied for the general equilibrium of the product and money markets in terms of the *IS* and *LM* functions. Now we study how these markets are brought into simultaneous equilibrium. It is only when the equilibrium pairs of interest rate and income of the *IS* curve equal the equilibrium pairs of interest rate and income of the *LM* curve that the general equilibrium is established. In other words, when there is a single pair of interest rate and income level in the product and money markets that the two markets are in equilibrium.

Such an equilibrium position is shown in Figure 6 where the *IS* and *LM* curves intersect each other at point *E* relating *Y* level of income to *R* interest rate. This pair of income level and interest rate leads to simultaneous equilibrium in the real or goods (saving-investment) market and the money (demand and supply of money) market. This general equilibrium position persists at a point of time, given the price level. If

there is any deviation from this equilibrium position, certain forces will act and react in such a manner that the equilibrium will be restored.

Consider point *A* on the *LM* curve where the money market is in equilibrium at Y_1 income level and R_2 interest rate. But the product market is not in equilibrium. In the product market, the interest rate R_2 is lower. The product market can be in equilibrium at Y_1 income level only at a higher interest rate R_1 corresponding to point *B* on the *IS* curve. Consequently at point *A*, there is excess of investment over saving since point *A* lies to the left of the *IS* curve. The excess of *I* over *S* indicates excess demand for goods which raises the level of income. As the level of income rises, the need for transactions purposes increases. In order to have more money for transactions purposes, people sell bonds. This tends to raise the interest rate. This moves the *LM* curve from point *A* upward to point *E* where a combination of higher interest rate *R* and higher income level *Y* exists. On the other hand, rising interest rate reduces investment and an increasing income raises saving. This helps to bring about the equality of *I* and *S* at point *E* where the general equilibrium is re-established by the equality of *IS* and *LM*.

Now consider point *C* on the *IS* curve in Figure 6 where the product market is in equilibrium at R_2 interest rate and Y_2 income level. The money market is not in equilibrium. It can be in equilibrium at Y_2 income level only at a higher interest rate R_1 corresponding to point *D* on the *LM* curve. At point *C*, the demand for money (*L*) is greater than the supply of money (*M*) because point *C* reflects lower rate of interest R_2 than is required for the equality of *L* and *M*. Thus there is excess demand for money at R_2 interest rate. The excess demand for money leads people to sell bonds but there is less demand for bonds which tends to raise the interest rate. When the rate of interest begins to rise, the product market is thrown into disequilibrium because investment falls. Falling investment leads to falling income which in turn reduces saving. This process

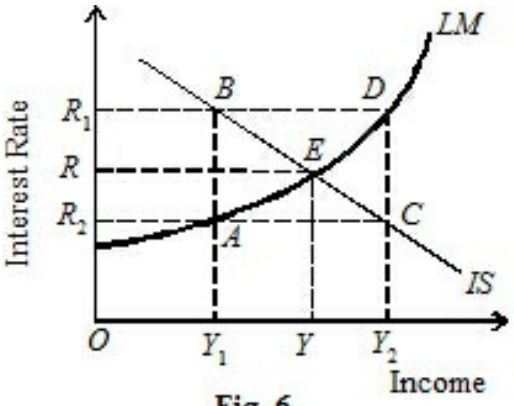


Fig. 6

ultimately brings the equilibrium of the product market when $I=S$ at point E. On the other hand, falling income reduces the transactions demand for money and ultimately brings about the equality of LM at point E where the equilibrium is re-established by the equality IS and LM curves, at R interest rate and Y income level.

CHANGES IN GENERAL EQUILIBRIUM

The general equilibrium of the product and money markets discussed above is based on the static equilibrium analysis. It started from a point of disequilibrium and again reached the equilibrium point of the equality of IS and LM functions. But the general equilibrium combination of Y income level and R rate of interest may change either due to a shift in the IS function or the LM function, or by both the functions shifting simultaneously. The IS function may shift due to changes in the saving function or the investment function. The shifts in the LM function may be caused by changes in the money supply or liquidity preference.

The shifting of the IS curve to the right and the consequent equilibrium with the given LM curve is illustrated in Figure 7. With the increase in the autonomous investment (or reduction in saving), the IS curve moves from IS to IS_1 and the new equilibrium is established at point E_1 which indicates a higher level of income Y_1 at a higher interest rate R_1. If the interest rate had not increased but remained at R level, the increse in investment would have raised income from Y to Y_2 level. But this much increase in income would not take place. This is because with the increase in income, the demand for money for transactions purposes will raise the interest rate to R_1 level where the IS and LM functions intersect at point E_1. Thus the expansionary effect of increase in investment is dampened by the rise in the interest rate and the income rises by less than the full multiplier.

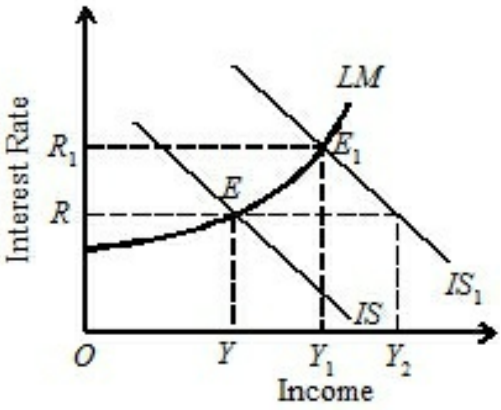


Fig. 7

In the opposite case when investment falls or saving increases, the *IS* function will shift to the left and the equilibrium will be established at a lower level of income and interest rate. This situation has not been depicted in Figure 7.

With the increase in the money supply, the *LM* curve shifts to the right as LM_1 which moves the economy to a new equilibrium point E_1 where the *IS* curve intersects the *LM* curve in Figure 8. The increase in the money supply brings down the interest rate R_1 in the money market. This, in turn, increases investment thereby raising the level of income to Y_1. Thus the effect of the increase in money supply is to shift the *LM* function to the right and a new equilibrium is established at a lower interest rate and higher income level. But how much income will rise as a result of an increase in the money supply depends on (1) how much the interest rate falls which in turn depends on the elasticity of speculative demand for money, and (2) how much investment rises as a result of any given fall in the interest rate which in turn depends on the interest-elasticity of investment demand function.

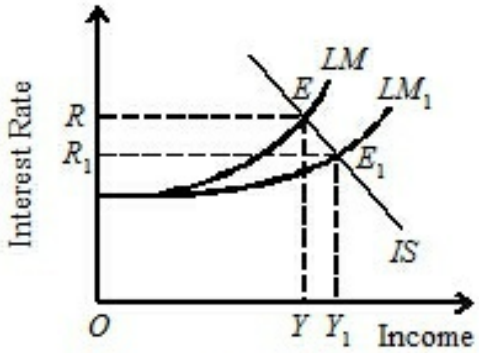


Fig. 8

Contrariwise, a decrease in the money supply or an increase in the demand for money will shift the *LM* function to the left such that a new equilibrium is established at a higher interest rate and lower income level. This case has not been depicted in Figure 8.

Simultaneous Shifts in the *IS* and *LM* functions

We have seen above that with the increase in investment when the *IS* curve shifts to the right, both the rate of interest and the level of the income tend to rise, given the *LM* curve. On the other hand, when an increase in money supply shifts the *LM* curve to the right, it lowers the rate of interest and raises the income level, given the *IS* curve.

Now suppose both the *IS* and *LM* curves shift to the right simultaneously

as a result of the increase in investment and money supply respectively. How will these *expansionary fiscal and monetary policies* affect the level of income and the rate of interest ? This is illustrated in Figure 9 where the increase in investment shifts the *IS* curve to IS_1 and the increase in the money supply shifts the *LM* curve to LM_1. Consequently, the new equilibrium position is E_1 where the IS_1 and LM_1 curves intersect. The rate of interest remains at the old level R_1 but the income increases from Y to Y_2. Given the money supply with no change in the *LM* curve, an increase in investment would raise both income and the rate of interest. This is shown in the figure when the IS_1 curve intersects the *LM* curve at E_2 and the interest rate rises to R_2 and income to Y_1. But the rise in income is slowed down because of the rise in the interest rate. If the money supply increases by so much as to prevent the rise in the interest rate, the increase in income will be equal to the full expansionary effect of the rise in investment. This is depicted in the figure by the shifting of the *LM* curve to the right as LM_1 which intersects the IS_1 curve at E_1. The income increases to Y_2 but the rate of interest remains at the same level R_1. So there has been full income-expansionary effect of the increase in investment as a result of the simultaneous increase in money supply by just the amount necessary to prevent the rise in the interest rate.

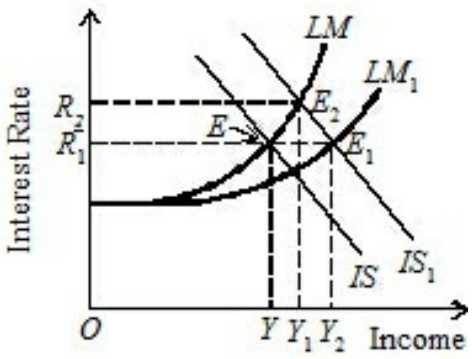


Fig. 9

Table I summarises the causes and directions of shifts in the *IS* and *LM* curves.

Table I : Causes and Directions of Shifts in IS and LM curves

Cause	*Curve*	*Direction*
Increase in investment or consumption	*IS*	Right
Decrease in investment or increase in savings	*IS*	left
Increase in money supply or decrease in money demand	*LM*	left

Decrease in money supply or increase in money demand	*LM*	right

EXERCISES

1. Use IS-LM framework to explain the joint determination of the rate of interest and the level of income.
2. Use the IS-LM framework to analyse the effects of an autonomous increase in investment.
3. Derive an IS and LM curve. Give their properties. [Hint: Explain their slopes and shifts.]
4. Use the IS-LM technique to analyse the effects of simultaneous increase in autonomous investment and money supply.

CHAPTER 52

EXTENSIONS OF IS-LM MODEL

INTRODUCTION

In the previous chapter, we studied the determination of equilibrium levels of national income and interest rate with the help of *IS-LM* curves. This chapter analyses the effects of changes in monetary and fiscal policies by the government by using the *IS-LM* curve model. We also extend the *IS-LM* model to study the effects of flexible prices and labour market on equilibrium levels of national income and interest rate.

EFFECTS OF CHANGES IN MONETARY AND FISCAL POLICIES BY THE GOVERNMENT

The equilibrium levels of national income and interest rate are shown by the intersection of the *IS* and *LM* curves. When the government changes its monetary policy or fiscal policy, either the *LM* curve or the *IS* curve shifts and the equilibrium levels also change. In the *IS-LM* model, monetary policy is represented by the *LM* curve and fiscal policy by the *IS* curve.

Effects of Changes in Monetary Policy

In the *IS-LM* model, monetary policy is represented by the *LM* curve. Suppose the government adopts an expansionary monetary policy to control deflation in the economy. For this, it increases the money supply through its central bank. The increase in the money supply is shown by shifting the *LM* curve to the right as LM_1 curve in Fig. 1. When the money

supply increases, the interest falls, given the price level. The fall in the interest rate increases investment demand which causes the income to rise. This in turn, increases consumption demand. The fall in the interest rate and the rise in income jointly increase the aggregate demand and national income. As a result, a new equilibrium is established in the *IS-LM* model with lower interest rate and higher income levels.

This is illustrated in Fig. 1 where the initial equilibrium point *E* is at interest rate or and income level *OY*. An increase in the money supply shifts in *LM* curve to the right to LM_1. It intersects the given *IS* curve at point E_1 which shows the new equilibrium with fall in the interest rate from *OR* to OR_1 and rise in the national income from *OY* to OY_1. On the other hand if the government wants to control inflation, it reduces the money supply which shifts the *LM* curve to the left. As a result, a new equilibrium point will be established at higher interest rate and lower national income level. This will be due to the effects of reduction in the money supply when the interest rate rises, aggregate demand falls and national income declines.

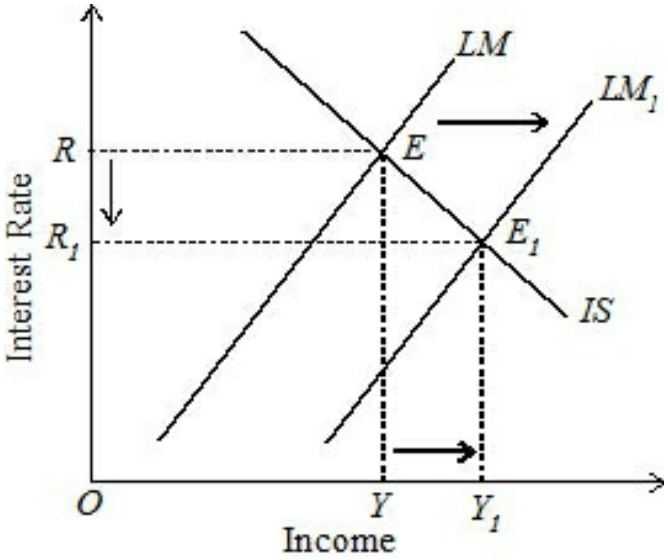


Fig. 1

EFFECTS OF CHANGES IN FISCAL POLICY

The effects of changes in fiscal policy are related to government expenditure and taxes which are shown by shifts in the *IS* curve. They are explained in the case of an expansionary fiscal policy.

1. Increase in Government Expenditure

Suppose there is depression in the economy and the government wants to raise the level of employment and income. For this, it increases its

expenditure which raises aggregate demand both directly as government demand rises and indirectly when consumer expenditure increases with rise in employment and income of the people. As income rises, the transactions demand for money increases. The money supply being fixed, the increase in transactions demand leads to reduction in the speculative (bonds) demand for money. This causes the interest rate to rise. Thus with the increase in public expenditure the equilibrium levels of income and interest rate rise.

This is illustrated in Fig. 2 where the initial equilibrium point is *E* at *OR* interest rate and *OY* income level. The increase in government expenditure shifts the *IS* curve to the right to IS_1 which intersects the given *LM* curve at point E_1. This results in rise in the interest rate from *OR* to OR_1 and of the national income from *OY* to OY_1. The figure shows that the horizontal distance by which the *IS* curve shifts when government expenditure increases

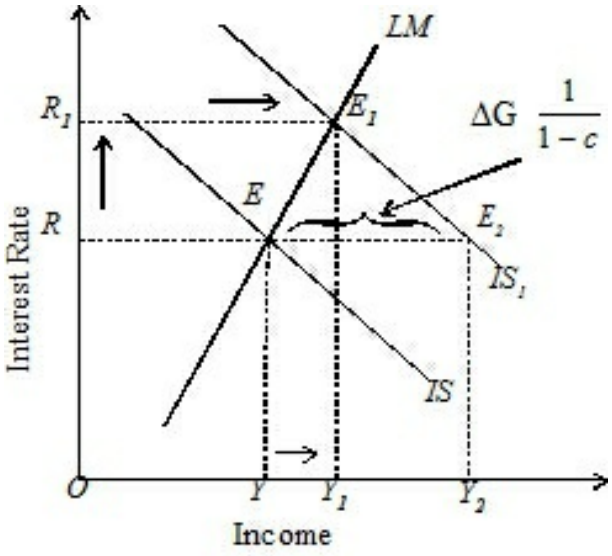


Fig. 2

is equal to $\Delta G \frac{1}{1-c}$, where ΔG is increase in governement expenditure and (1/1–*c*) is the multiplier in the Keynesian model. This leads to increase in equilibrium income from *OY* to OY_2. But in the *IS-LM* model, income rises to OY_1 which is less than OY_2. This is because in the *IS-LM* model when the interest rate rises with the increase in government expenditure, it causes *crowding out* (decline) in some private investment.

The opposite will be the effects of a decrease in government expenditure when there is inflation in the economy and the government adopts contractionary fiscal policy.

2. Reduction in Taxes

We now explain the effects of reduction in taxes in the case of

expansionary fiscal policy. A reduction in taxes raises the disposable income and increases consumption of the people. As income increases, the demand for money also rises and the demand for bonds declines. This leads to rise in interest rate. Thus the equilibrium levels of income and interest rate rise.

Figure 3 illustrates that a reduction in taxes ($-\Delta T$) shifts the *IS* curve to the right to IS_1. Income increases from *OY* to OY_1 and the interest rate rises from *OR* to OR_1. The figure also shows that the horizontal distance by which the *IS* curve shifts with increase in taxes which is equal to $-\Delta T \left[\frac{c}{1-c}\right]$, is the tax multiplier in the Keynesian model. This leads to the rise in income by EE_2 ($=YY_2$) at the initial interest rate *OR*. But in the *IS-LM* model, the cut in taxes causes the interest rate to rise to OR_1 which reduces investment. As a result, the rise in income by YY_1 is less than YY_2. This is because in the Keynesian model, investment is assumed to be fixed.

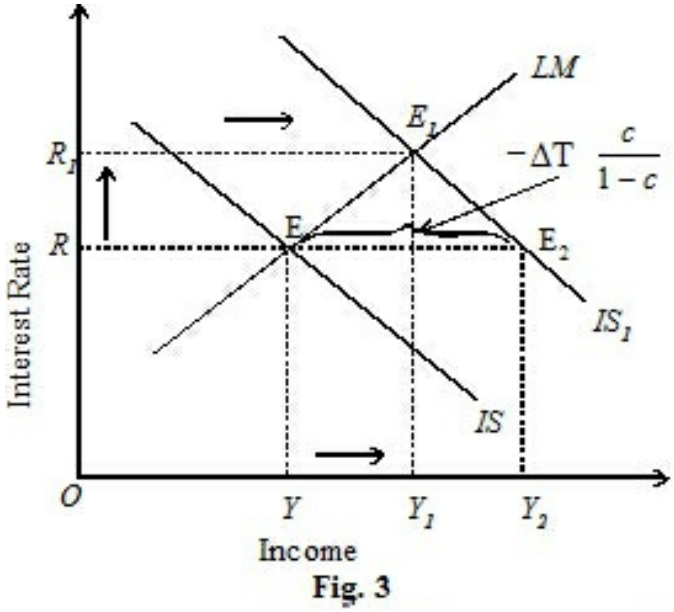


Fig. 3

In the opposite case of increase in taxes both income and interest rate will decline in contractionary fiscal policy.

MONETARY-FISCAL POLICY MIX

We have been above that both monetary and fiscal policies affect income. But their effects on the interest rate and investment are different. When expansionary monetary policy is adopted, the interest rate declines and investment increases. But in expansionary fiscal policy, when government expenditure is increased or taxes are cut, interest rate rises and investment declines. To keep the interest rate low and to encourage investment, the

government adopts a monetary-fiscal mix of an *accommodating monetary policy* alongwith an expansionary fiscal policy in which the increase in money supply will prevent the interest rate from rising and thus offset the crowding out of private investment. This is illustrated in Fig. 4 where the initial equilibrium is at point *E* where the *IS* and *LM* curves intersect and determine *OR* interest rate and *OY* income level. With the increase in government expenditure or tax cut, the *IS* curve shifts to the right as IS_1 curve. It cuts the *LM* curve at point E_1 and the new equilibrium interest rate is OR_1 and income level is OY_1. But the increase in interest rate leads to the crowding out of some private investment. To prevent this crowding out, the government adopts an accommodating monetary policy in which the money supply is increased sufficiently so that the *LM* curve shifts far enough to the right to LM_1 curve. It cuts the IS_1 curve at point E_2, the interest rate remains at the original level OR but income rises to OY_2. Thus this monetary-fiscal mix has raised the income level with the interest rate remining or *OR* level.

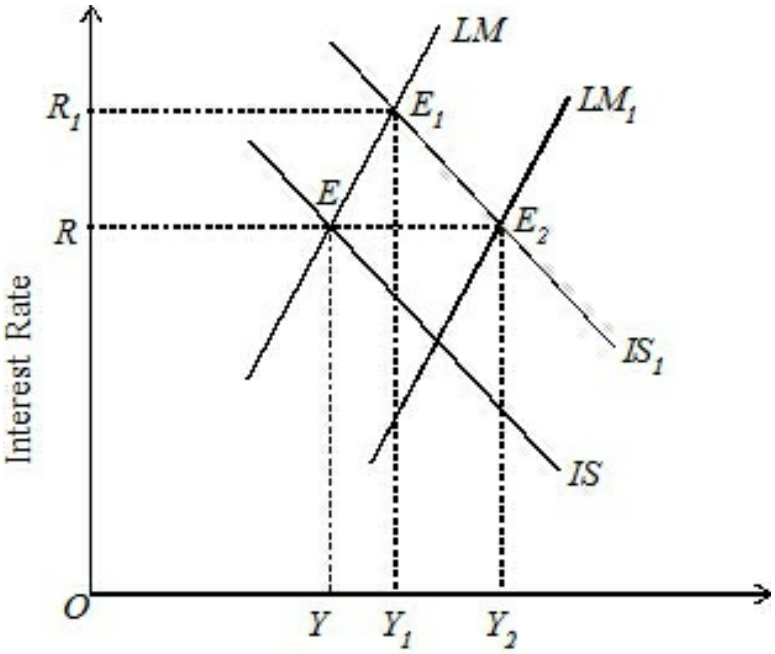


Fig. 4

IS-LM Model with Labour Market and Flexible Prices

The *IS-LM* model with labour market and flexible prices explains the determination of interest rate, price level and employment, output and income. It is a *three-sector model* in which goods, money and labour markets are in equilibrium. The entire analysis presents a synthesis of the Keynesian and classical systems based on the interaction of the aggregate

demand and aggregate supply curves which are derived from the *IS* and *LM* curves.

The Keynesian system is based on demand-determined output in which prices and wages are fixed, given the usual upward sloping supply curve to the right. On the other side, the neo-classical system is based on supply-determined output in which prices and wages are *flexible*, given a vertical supply curve.

In the analysis that follows, the aggregate demand (*AD*) curve is derived from the *IS* and *LM* curves. The *AD* curve is common to both the neo-classical and Keynesian systems of the model. The main difference lies in the shapes of the neo-classical and Keynesian aggregate supply (*AS*) curves. The derivation of the new-classical and Keynesian *AS* curves from the *IS* and *LM* curves is not possible. So the respective shapes of the *AS* curves are taken alongwith the *AD* curve to explain the two systems. Finally, the upward sloping *AS* curve to the right is taken to analyse the synthesis of the two systems in terms of the *IS* and *LM* model.

The Aggregate Demand Curve (AD)

The *AD* curve is drawn by plotting each equilibrium level of income (output) that is associated with each price level. All points on the *AD* curve represent the equilibrium of the product (goods) market and the money market.

In this analysis, all vriables like investment, government spending, saving and taxes of the product market are assumed fixed and are, therefore, not affected by a change in the price level. So a change in the price level does not shift the IS curve. In the case of the *LM* curve, the variables are the real money supply and a constant money demand curve. So with the change in the price level, the real money supply (*M/P*) changes which shifts the *LM* curve and produces a new equilibrium level of income with the *IS* curve. Plotting income against the given price level, gives one point of the *AD* curve.

The derivation of the *AD* curve is shown in Fig. 5 where in Panel (A) the initial equilibrium is at point E_2 with income *OY* and interest rate OR_2

when the *IS* curve intersects the LM_2 curve. Plotting *Y* against the price level OP_2 in Panel (A) of the figure gives the point *A*. Thus the income level at which the *IS* and *LM* curves intersect for a given price level is a point on the *AD* curve. *A* fall in the price level to OP_1, shifts the LM_2 curve to the right to LM_1 in Panel (A). The new equilibrium is at E_1 with income OY_1 and and interest rate OR_1. Plotting Y_1 against P_1 in Panel (B), gives the point *B*. A further fall in the price to *OP* shifts the LM_1 curve to the right to *LM* which increases equilibrium income to OY_2. Plotting Y_2 against *P* in Panel (B) yields the point *C*. By joining the points *A*, *B* and *C*, we trace the aggregate demand curve *AD*.

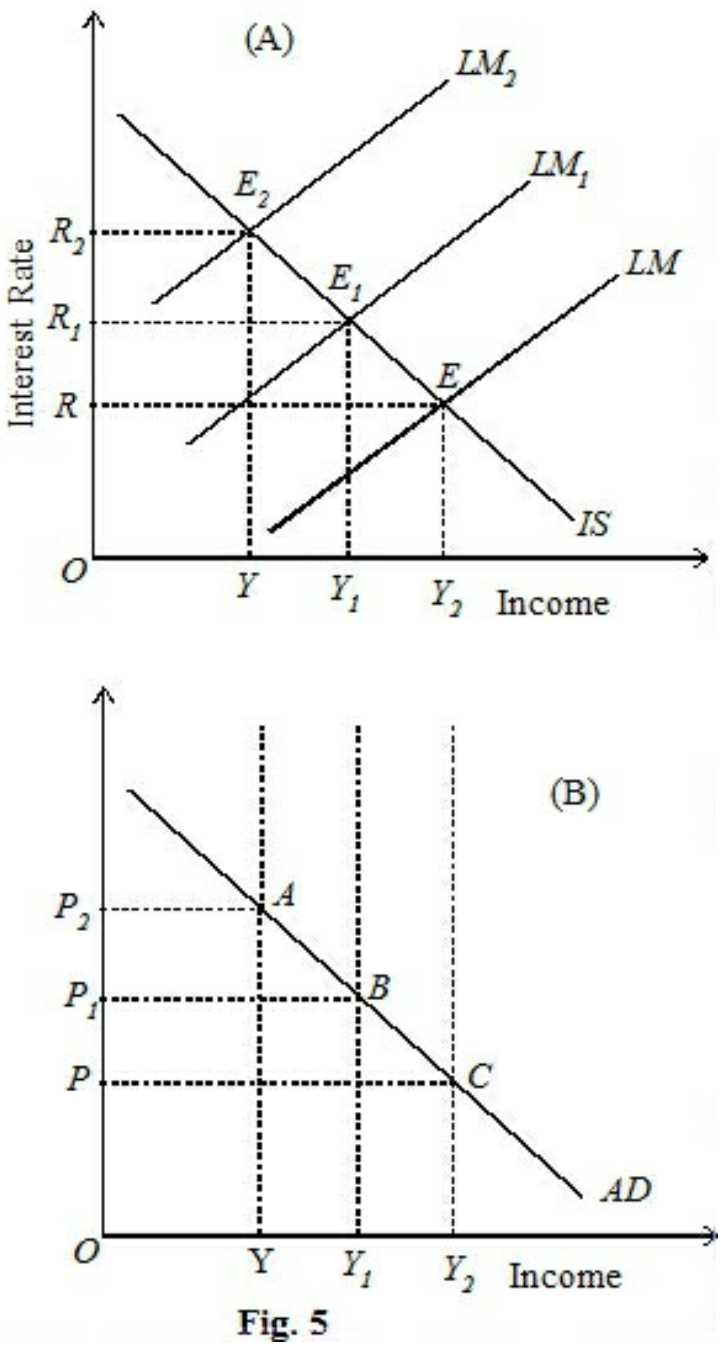


Fig. 5

The *AD* curve is negatively sloped because a fall in the price level increases the real money supply which leads to an excess supply of real money balances. With excess money supply, people buy bonds which raise bond prices and reduce interest rate. As a result of reduction in interest rate, investment increases which, in turn, increases output and income.

IS-LM Model with Flexible Wages and Prices : The Neo-Classical Analysis

We explain the *IS-LM* model with flexible wages and prices based on the neo-classical theory of employment, income and output. In this analysis, the Keynesian aggregate demand curve is combined with the neo-classical aggregate supply curve.

The Aggregate Supply Curve (AS)

The neo-classical *AS* curve is vertical as shown in Panel (C) of Fig.6 because of the following assumptions:

1. Wages and prices are fully flexible.

2. There is perfect information about market prices on the part of market participants. Both the employers of labour and workers know the money wage rate in the labour market and how much commodities the real wage rate (*W/P*) can buy.

3. The economy is always at its full employment income and output level.

4. The labour market is characterised by market-clearing which means that the equilibrium real wage is established at the full employment level.

5. Monetary and fiscal policy affect prices but not the aggregate level of output and employment.

Given these assumptions, the *AS* curve is derived from the demand for labour and supply of labour curves along with the aggregate demand function, as shown in Fig. 6.

Panel (A) of the figure shows labour market equilibrium at point *E* where the supply of labour curve, S_L equals the demand curve for labour curve, D_L which establish the equilibrium real wage rate *W/P*. This leads to the full employment equilibrium level, ON_F. As prices and wages are fully flexible, the real wage rate is also fully flexible. This means that the

labour market is always in equilibrium. Any excess supply of labour or demand for labour will instantaneously bid the real wage rate back to the equilibrium level. The aggregate production function. $Y=f(N,K,T)$ in Panel (B) shows the level of output OY at the full employment level ON_F. Since the level of full employment is unaffected by changes in prices, therefore the level of output is also independent of price changes. Consequently, the AS is vertical at the level of full-employment output OY_F in Panel (C) of the figure. To show the adjustment mechanism beginning from labour market equilibrium in Panel (A), a fall in the price level from OP to OP_1 with a given money wage rate would increase the real wage rate from W/P to W/P_1 . This would reduce the demand for labour thereby causing excess supply of labour by ds. The unemployed workers would bid down the money wage rate until the real wage rate is restored to its equilibrium level, W/P and the excess supply of labour is eliminated. Thus any change in the price level is immediately reflected in a change in money wage rate to maintain equilibrium real wage rate, full employment and output.

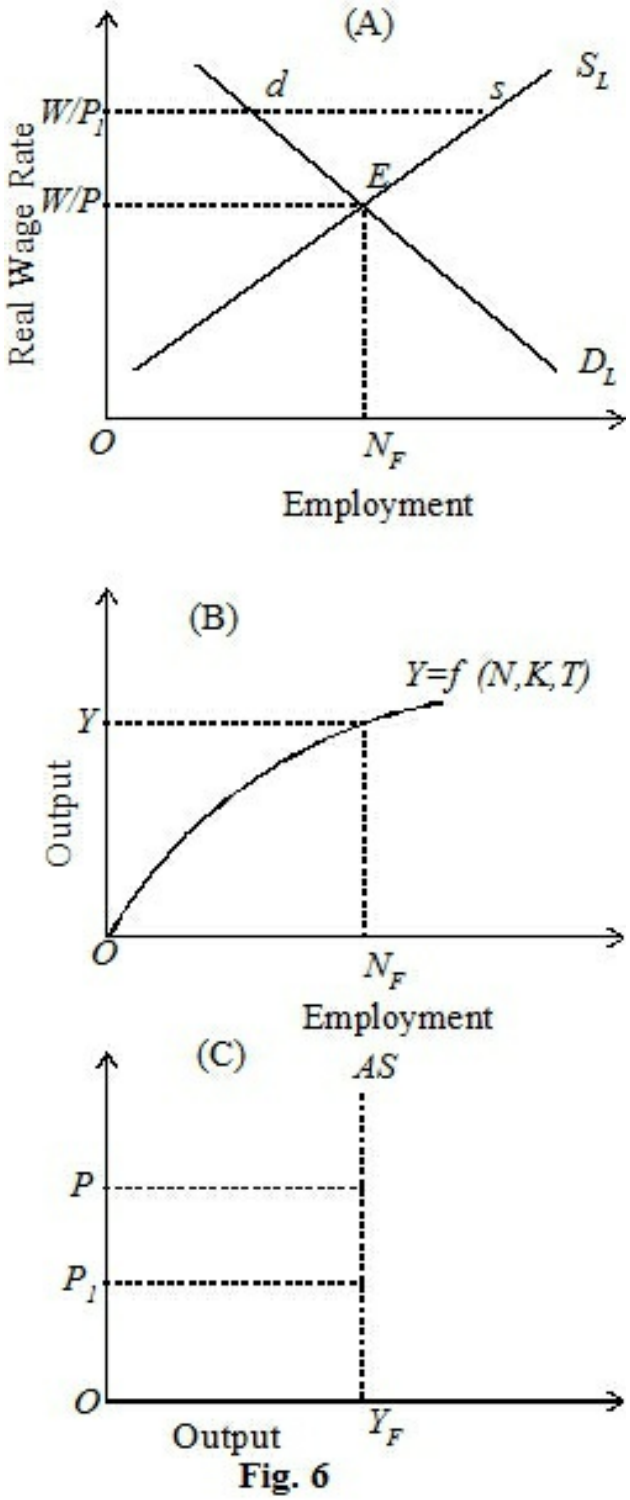


Fig. 6

The Complete Neo-Classical Model

Figure 7 depicts full equilibrium in the neo-classical model at point *E* where the *AD* curve intersects the *AS* curve. At this point, output is at its full employment level OY_F and the equilibrium price level is *OP*. With flexible money wages and prices, the real wage rate always adjusts to maintain full employment in the labour market. Given this labour, firms produce full employment output OY_F. The aggregate demand curve *AD* shows points at which money demand equals money supply and planned spending on goods equals output produced. The equilibrium price level *OP* clears the markets for labour, goods and money. The labour market is in equilibrium on the *AS* curve. The money market clears on the *AD* curve alongwith the equality of aggregate demand for goods with the actual output of goods. Thus in the neo-classical model, the markets for labour, money and goods are simultaneously in equilibrium at point *E*.

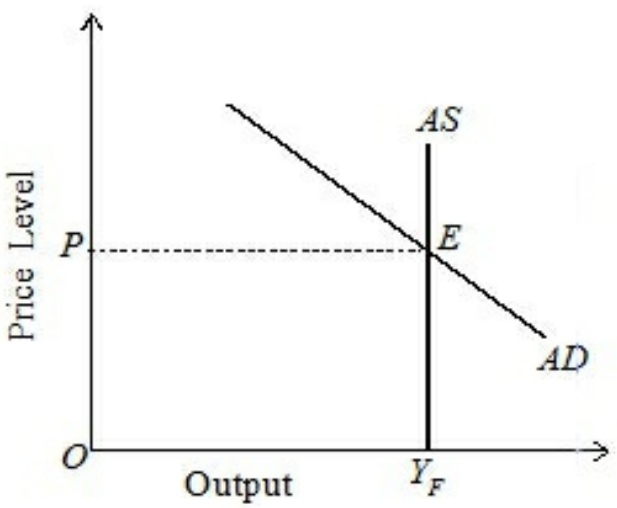


Fig. 7

Effect of Monetary Policy in the Neo-Classical Model

The effect of monetary policy on employment and output in the neo-classical system is explained in terms of the *IS-LM* model in Fig. 8. Suppose the economy is in equilibrium at point *E* where the *LM* curve intersects the *IS* curve with OR_1 interest rate and OY_F full employment income level in Panel (A). Panel (B) of the figure shows the equality of the *AD* and *AS* curves at point. *A* which determines OP_1 price with OY_F aggregate output in the economy.

An increase in the money supply shifts the LM_1 curve to the right to LM_2. The new equilibrium point is E_1 with lower interest rate OR_2 and higher income OY_1. The increase in the money supply and reduction in interest rate increase the aggregate demand which shifts the *AD* curve to the right to AD_1 in Panel (B). At the initial price level OP_1, the output increases to

OY_1. But point C on the AD_1 curve does not show the equilibrium price level which is at point B where the AD_1 curve intersects the vertical AS curve at the higher price level OP_2. The rise in the price level reduces the real money supply (M/P) which shifts the LM_2 curve to the left to LM_1 and the interest rate rises to OR_1 and the full employment equilibrium level of income and output is established at OY_F.

Let us understand the process through which the economy adjusts instantaneously in the neo-classical model from point E to E_1 in Panel (A) when the money supply is increased, and the price level is OP_1 in Panel (B). An increase in the money supply shifts the LM_1 curve to LM_2 and lowers the interest rate from OR_1 to OR_2. This increases the aggregate demand which shifts the AD curve to AD_1. But aggregate demand exceeds full employment output by (Y_F–Y_1) which the firms can supply at the initial price level OP_1. The excess demand for goods bids the price level to OP_2. High price level reduces the real

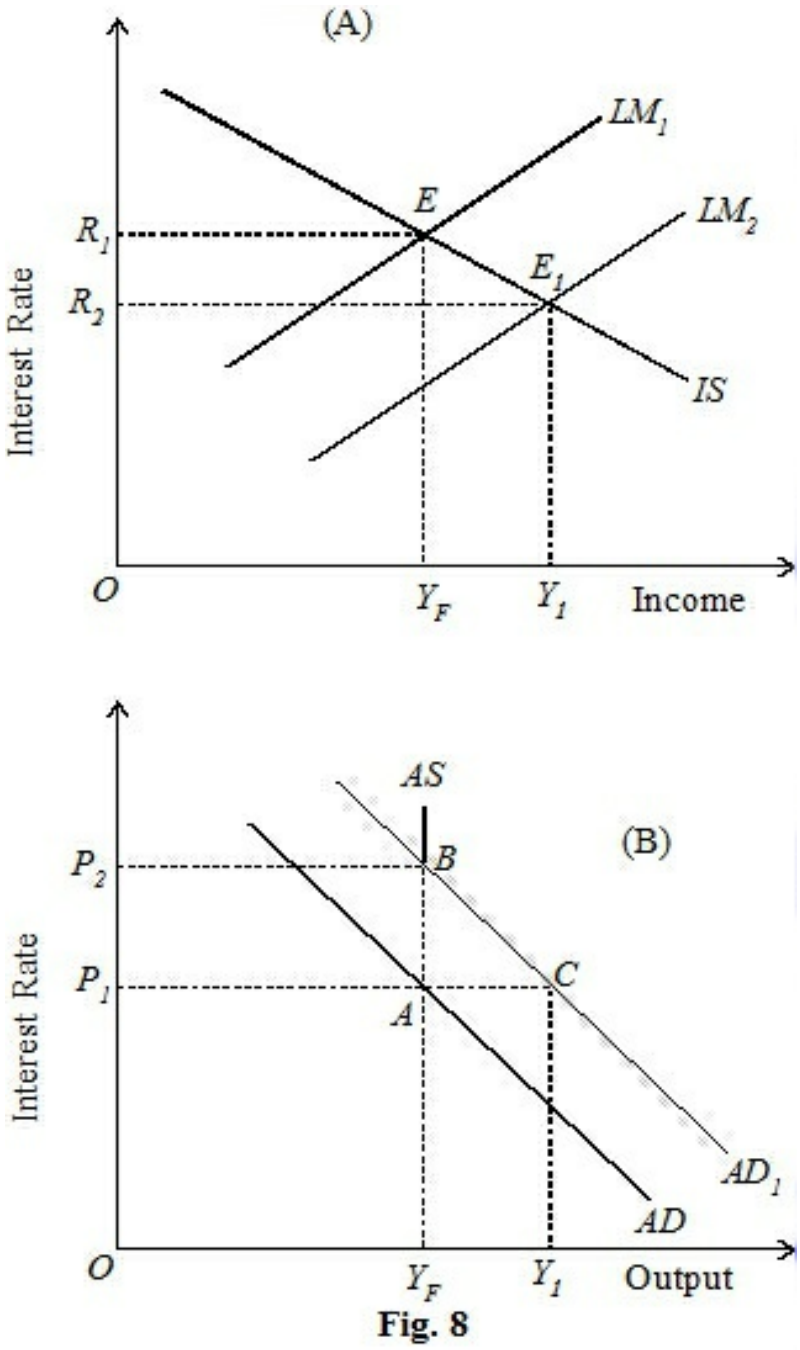


Fig. 8

money supply. As a result the *LM* curve and the interest rate return to their original levels LM_1 and OR_1 respectively and the economy is again at the full- employment income and output level OY_F.

In the labour market, high money wage rate matches the increase in the price level, thereby maintaining the real wage rate at its original level. Thus in the neo-classical model with full wage-price flexibility, a change in the money supply leads to an equal proportionate change in the money wage rate and the price level but without any change in income and output which remain at the full employment level.

Effect of Fiscal Policy in the Neo-Classical Model

Fig. 9 explains the effect of fiscal policy in the neo-classical model. An expansionary fiscal policy in the form of increase in government spending (or reduction in taxes) shifts the *IS* curve upward to the right from IS_1 to IS_2, given the *LM* curve, as shown in Panel (A). The effect of fiscal expansion is to increase aggregate demand which is shown by the shift in the *AD* curve upward to the right to AD_1 in Panel (B). At the initial price level OP_1, output increases to OY_1, more than the full employment output OY_F. But firms want to supply the full employment output OY_F. So there is excess demand for goods by *AC* (Y_F–Y_1). This bids up the price level until excess demand is eliminated at point *B* on the vertical *AS* curve at the higher price level OP_2. Given the money supply, the rise in the price level reduces the real money supply, raises the interest rate to OR_2 and reduces private expenditure on consumption and investment. Thus when an increase in government spending crowds out an equal amount of private expenditure, the IS_2 curve shifts leftward to IS_1 and the initial equilibrium level *E* is re-established so that the aggregate demand remains at the full employment level OY_F at the higher price level OP_2.

IS-LM Model in the Keynesian Analysis with Flexible Prices and Fixed Money Wages

The Keynesian model assumes that the money wage rate is not flexible

rather it is sticky downward in the short run. Therefore, workers are willing to accept a cut in their real wage rate by an increase in the general price level. Keynes argued that workers are prepared to work at the current money wage rate, even if their real wage rate is lowered by the increase in the price level. This is due to the existence of "money illusion" : they attach more importance to their money wage rate than to their real wage rate. In fact, they resist cut in their money wage rate but do not have the same resistance to a cut in real wage rate resulting from increase in the price level.

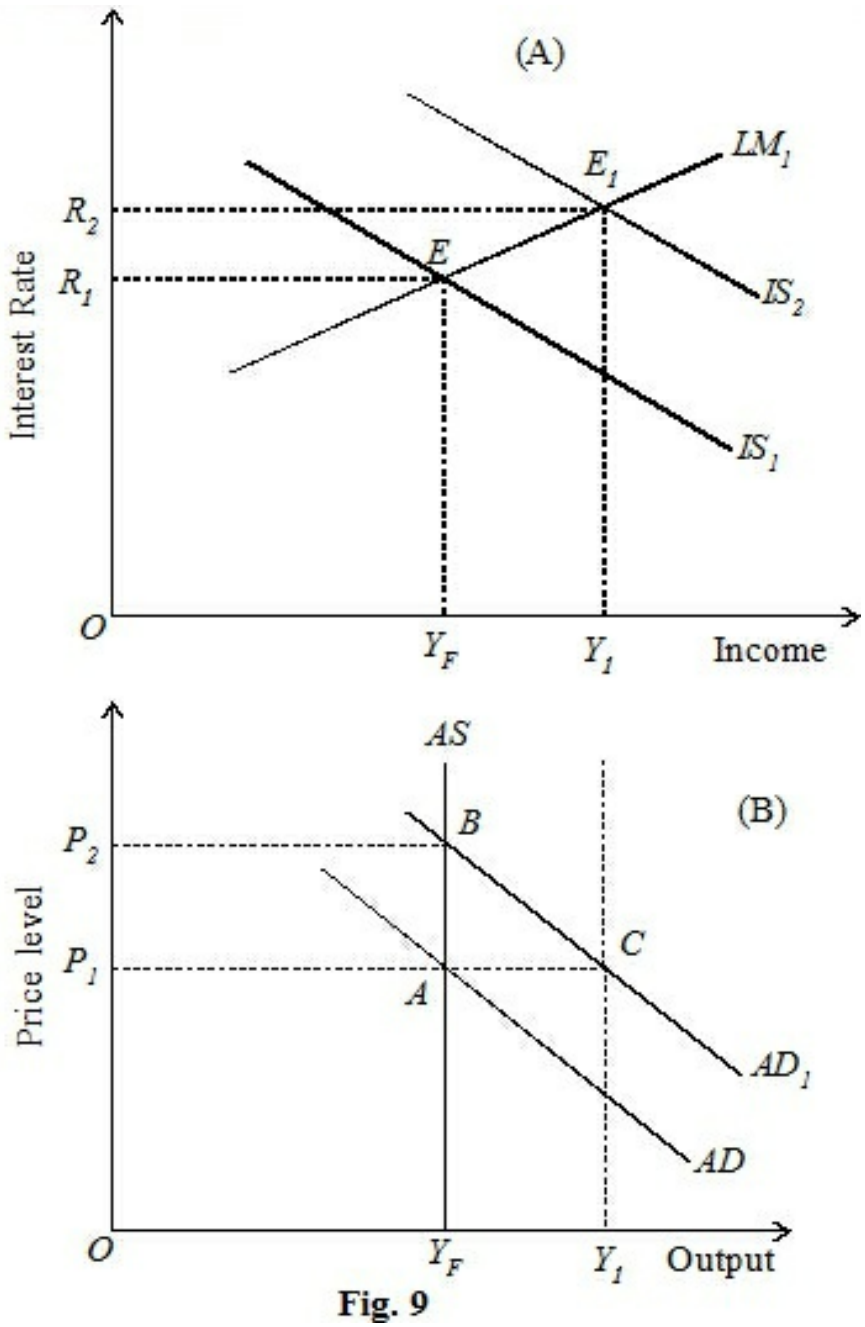


Fig. 9

Before discussing the complete Keynesian approach in terms of the *IS-LM* framework, it is essential to derive the Keynesian aggregate supply curve.

The Aggregate Supply Curve

Given a fixed money wage rate and flexible prices, the derivation of the aggregate supply curve, *AS* is shown in Fig. 10. Panel (A) of the figure depicts the determination of labour market equilibrium at point *E* when the demand for labour curve D_L intersects the supply of labour curve S_L at the market-clearing equilibrium real wage rate W/P_2. Starting movements of the real wage rate above point *E* and along the upper side of the D_L curve, when the price level is *OP* in Panel (C), the real wage corresponding to it is *W*/*P*. At this wage rate, *ON* workers are employed and the output supplied via the production function *f* (*N*,*K*,*T*) shown in Panel (B) is *OY*. As the price level rises to OP_1, the real wage rate falls to W/P_1,

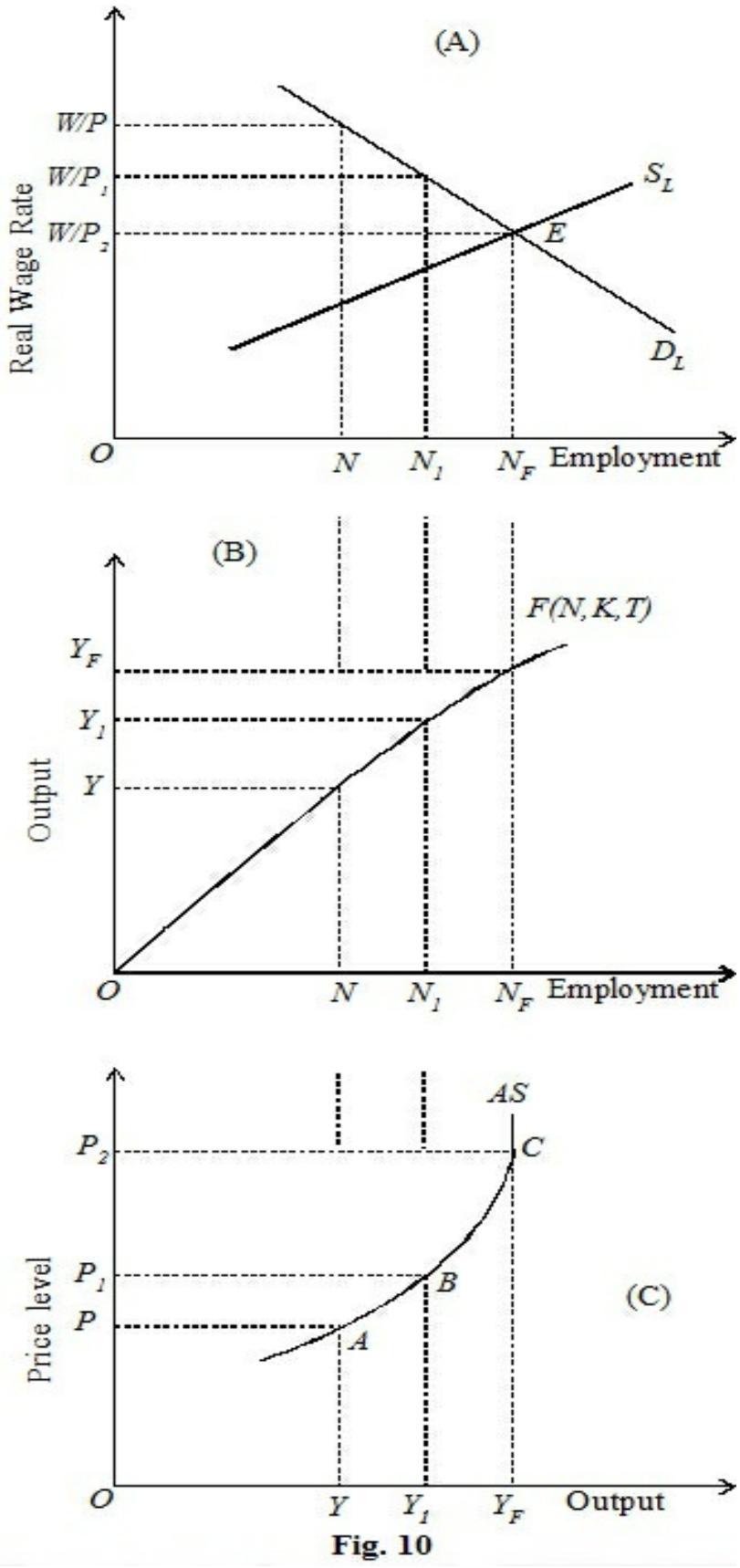


Fig. 10

thereby increasing employment to ON_1 due to money-illusion, and output to OY_1. Further rise in the price level to OP_2, increases employment to the market-clearing level ON_F and output to full employment level OY_F. The curve shown in Panel (C) which plots the price levels against the corresponding output levels at *A*, *B* and *C* trace the aggregate supply curve *AS*. The Keynesian *AS* curve which slopes upward from left the right is shown vertical beyond point *C* like the neo-classical *AS* curve. This is because with the rise in the price level, the output continues to rise up to the full employment level which is consistent with labour market clearing. But beyound this level, any further rise in the price level will have no effect on employment and output.

Thus the Keynesian *AS* curve represents a synthesis of the neo-classical supply curve and Keynes' aggregate supply curve. In the shortrun, the Keynesian *AS* curve is upward sloping like an ordinary supply curve because there is always underemployment in the economy. But in the longrun, the *AS* curve becomes vertical (from point C as in Fig. 11) when there is full employment in the economy.

The Complete Keynesian Model

The Keynesian system is presented in Fig. 11 where the *AD* curve (derived in Fig. 5) and the *AS* curve (derived in Fig. 10) determine the general price level and aggregate employment, output and income when the economy is in general equilibrium. In the Keynesian system, as money wages are constant, involumtary unemployment exists. Therefore, unlike the neo-classical *AS* curve which is vertical, the Keynesian *AS* curve is not vertical

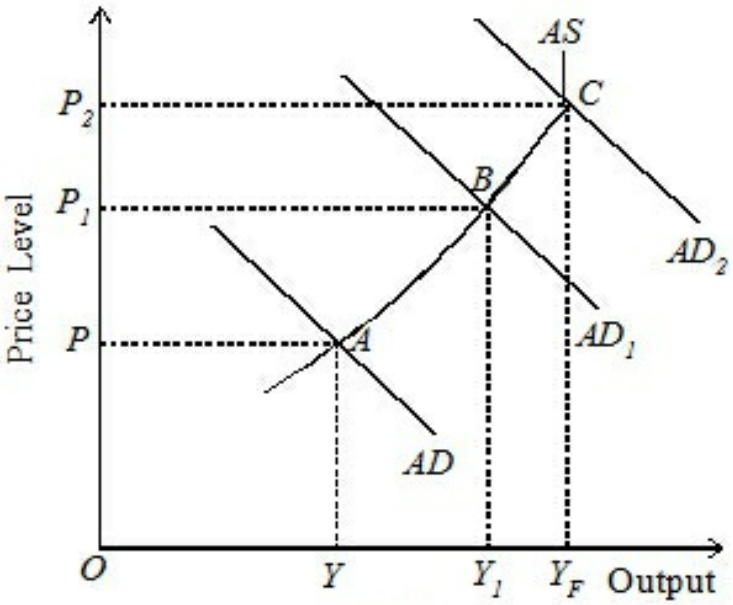


Fig. 11

until the full employment level is reached. To increase employment, the Keynesians suggest increase in aggregate demand which affects the price level and aggregate output and income. In Fig. 11 equilibrium occurs at point *A* when the *AS* and *AD*curves intersect at the price level *OP* and output level *OY*. When the *AD* curve shifts upward to AD_1, the price level rises to OP_1 and output level to OY_1. At this level, the actual output falls short of full employment output by $Y_1–Y_F$. If aggregate demand is increased sufficiently so that the AD_1 curve shifts to AD_2 curve which intersects the *AS* curve at point *C*, the price level would increase from OP_1 to OP_2 and the aggregate output would increase from OY_1 to the full employment level OY_F. Any further increase in aggregate demand would have no effect on aggregate output, employment and income, except raising the price level. This is because the demand for productive resources would exceed their available supplies at full employment.

In contrast, in the neo-classical system, the *AS* curve is a vertical straight line (see Fig. 3) which alongwith the Keynesian aggregate demand curve, *AD* determines only the equilibrium price level, *OP* and the level of full employment output OY_F. Any increase in aggregate demand would shift the *AD* curve upward to the right, thereby causing only increase in the price level at the same level of output OY_F.

Effect of Monetary Policy in the Keynesian System

The effect of monetary policy on employment, output and income in the Keynesian system is explained in Fig. 12 in terms of the IS-LM model when the price level is flexible and money wage rate is fixed. Suppose the economy is in equilibrium at point *E* where the *LM* curve intersects the *IS* curve with *OR* interest rate and *OY* employment and income level in Panel (A) of the figure. The equality of the *AD* curve and *AS* curve is depicted at point *A* in Panel (B) which determines *OP* price with *OY* aggregate output. An increase in the money supply to achieve full employment shifts the *LM* curve to the right to LM_2 which cuts the *IS* curve at point E_2 which leads to the full employment level of income, OY_F. This shifts the *AD* curve to the right to AD_2 and raises the price level from *OP* to OP_2. The increase in

the price level reduces the real money supply (M/P) which partially offsets the effects of the rise in the money supply so that the LM_2 curve shifts to the left to LM_1 and cuts the IS curve at E_1. Now the interest rate rises from OR_2 to OR_1 and the level of employment and income is OY_1 which is less than the full employment level OY_F. With the rise in interest rate, investment declines and the aggregate demand curve AD_2 shifts leftward to AD_1 with fall in the price level from OP_2 to OP_1 and output from OY_F to OY_1.

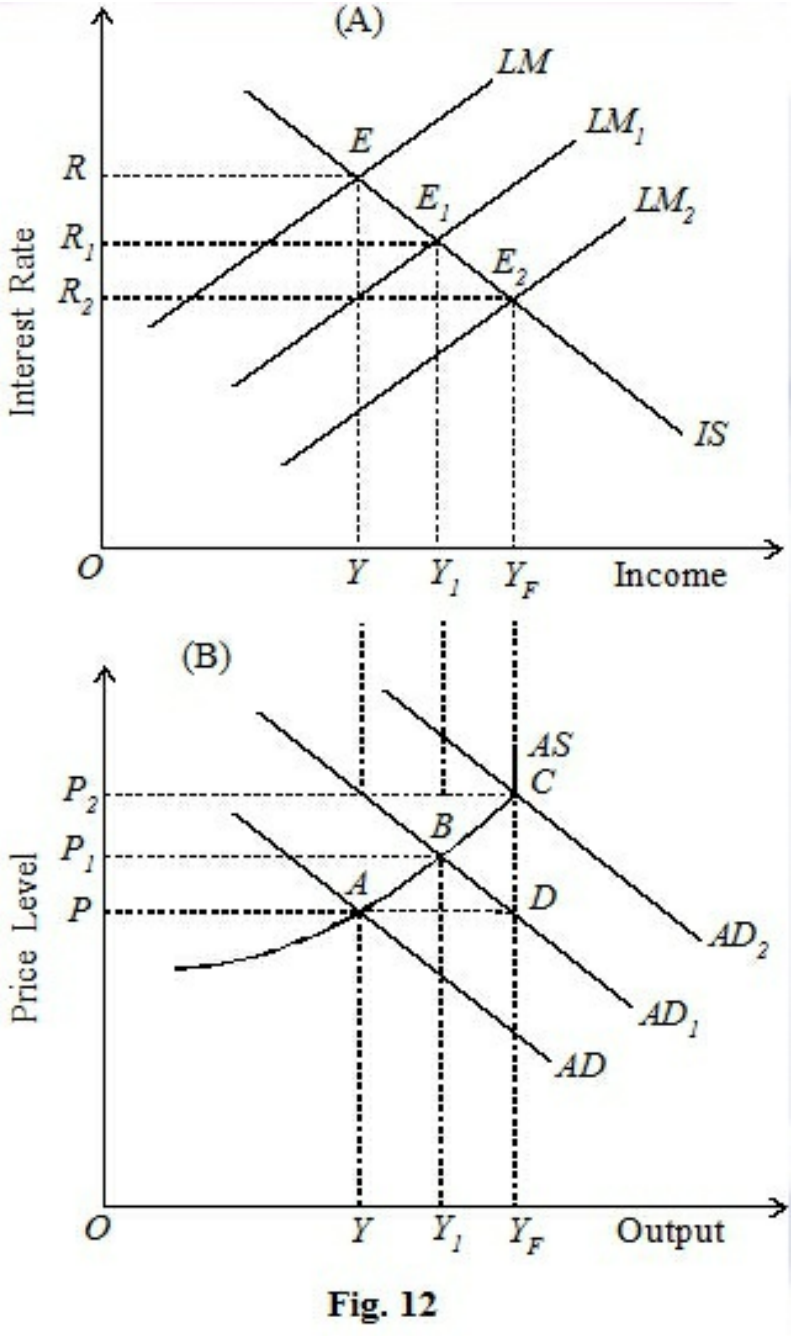


Fig. 12

However, the full employment level of output OY_F can also be achieved, if the price level remains constant at OP when the AD_1 curve intersects the AS curve at point D. But this is not possible under flexible prices and fixed money wage rate in the Keynesian system, because there is always underemployment equilibrium and monetary policy is less effective.

Effect of Fiscal Policy in the Keynesian Model

The effect of fiscal policy on employment, income and output in the Keynesian system is explained in Fig. 13 in terms of the *IS-LM* model when the price level is flexible and money wage rate is fixed. Suppose the government follows an expansionary fiscal policy by increasing investment or reduction in taxes to attain full employment in the economy. Initially, the economy is in equilibrium at point *E* where the *IS* curve cuts the *LM* curve with *OR* interest rate, *OY* income level and *OP* price level. As a result of increase in investment, the *IS* curve shifts upward to the right to IS_1 which intersects the *LM* curve at point E_2 in Panel (A) of the figure. This raises the interest rate from *OR* to OR_2 and income from *OY* to the full employment level OY_F. The increase in investment is reflected in the shift of the aggregte demand curve from *AD* to AD_2 where it cuts the *AS* curve at point *C* and the price level rises from *OP* in Panel (B) to OP_2 and output from *OY* to OY_F. The rise in the price level decreases the real money supply which shifts the *LM* curve upward

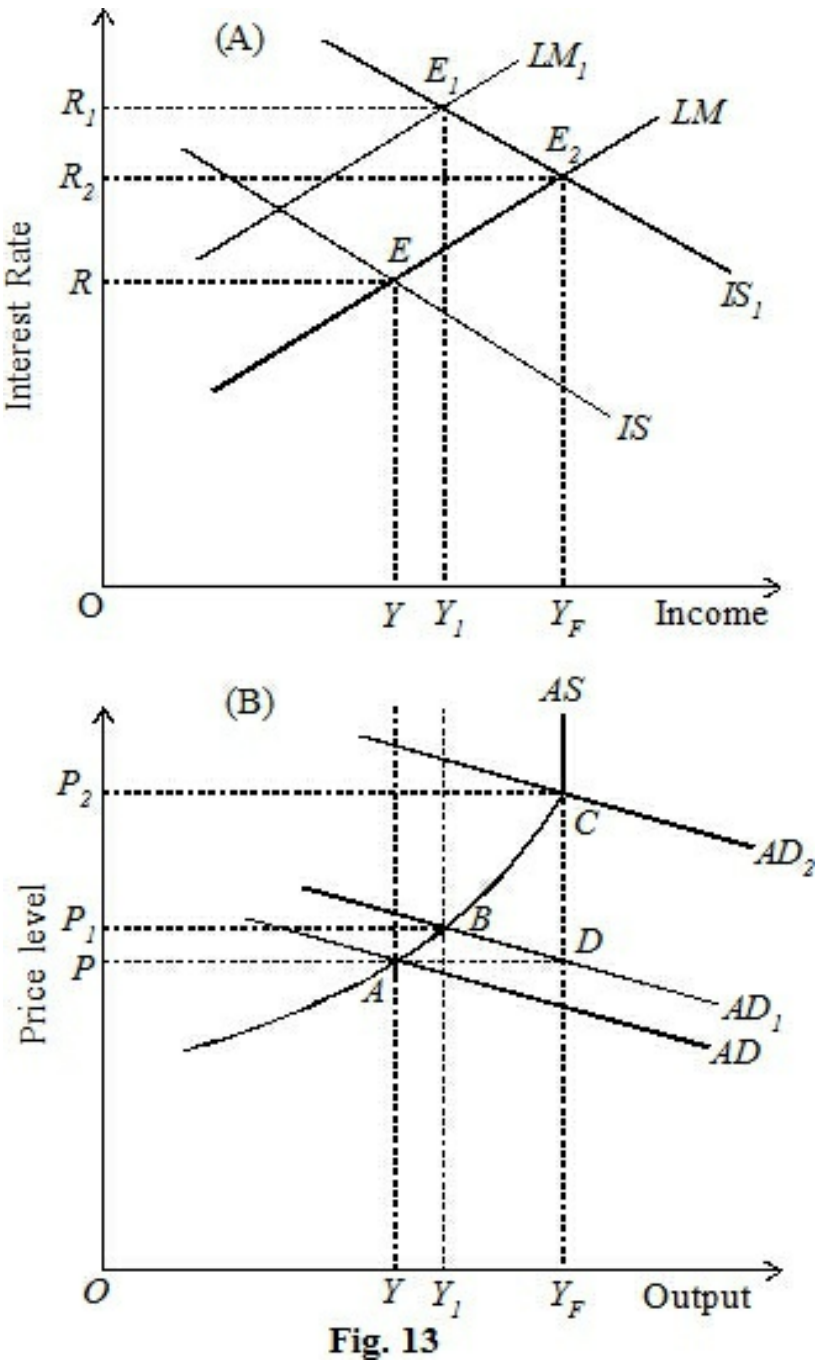


Fig. 13

to the left to LM_1 and cuts the IS_1 curve at point E_1 thereby raising the interest rate from OR_2 to OR_1 and reducing the income level from OY_F to OY_1. The rise in the interest rate leads to the decline in aggregate demand which shifts the AD_2 curve downward to AD_1 which cuts the *AS* curve at point *B*. The price level falls from OP_2 to OP_1 and and output from OY_F to OY_1.

However, the full employment output level, OY_F can also be achieved if the price level is constant at *OP* and the AD_1 curve intesects the *AS* curve at point *D*, as shown in Panel (B) of the figure.

But this is not possible with flexible prices and fixed money wage rate in the Keynesian system because there is always underemployment equilibrium, and fiscal policy, like monetary policy, is less effective.

SELECTED READINGS

R.J. Gordon, Macroeconomics, 6/e, 1993.

R.E. Hall and J.B. Taylor, Macroeconomics, 5/e, 1997

N.G. Mankiw, Macroeconomics, 4/e, 2004.

R.T. Froyen, Macroeconomics, 7/e, 2003

EXERCISES

1. Explain the effects of government monetary and fiscal policies within the IS-LM curve model.
2. Explain within the IS-LM curve model the effects of monetary and fiscal policies when wages and prices are flexible.
3. Explain within the IS-LM curve model the effects of monetary and fiscal policies when prices are flexible and wages are fixed.
4. Explain why the *AD* curve slopes downward in the IS-LM curve model.

5. Explain why the *AS* cure slopes upward in the IS-LM curve model.

CHAPTER 53

EFFECTIVENESS OF MONETARY AND FISCAL POLICY

INTRODUCTION

The relative effectiveness of monetary and fiscal policy has been the subject of controversy among economists. The monetarists regard monetary policy more effective than fiscal policy for economic stabilisation. On the other hand, the Keynesians hold the opposite view. In between these two extreme views are the synthesists who advocate the middle path. Before we discuss them, we study the effectiveness of monetary and fiscal policy in terms of shape of the *IS* curve and the *LM* curve. The *IS* curve represents fiscal policy and the *LM* curve monetary policy.

MONETARY POLICY

The government influences investment, employment, output and income through monetary policy. This is done by increasing or decreasing the money supply by the monetary authority. When the money supply is increased, it is an expansionary monetary policy. This is shown by shifting the *LM* curve to the right. When the money supply is decreased, it is a contractionary monetary policy. This is shown by shifting the *LM* curve to the left.

Figure 1 illustrates an expansionary monetary policy with given *LM* and *IS* curves. Suppose the economy is in equilibrium at point *E* with *OY*

income and *OR* interest rate. An increase in the money supply by the monetary authority shifts the *LM* curve to the right to LM_1, given the *IS* curve. This reduces the interest rate from *OR* to OR_1 thereby increasing investment and national income. Thus the national income rises from *OY* to OY_1.

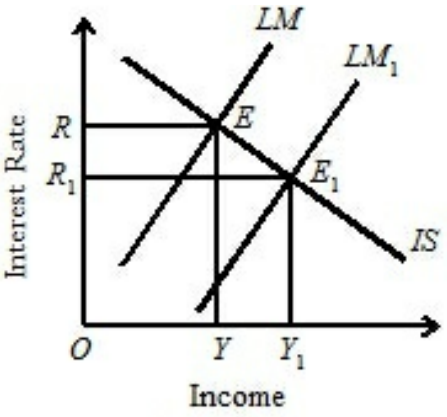


Fig. 1

But the

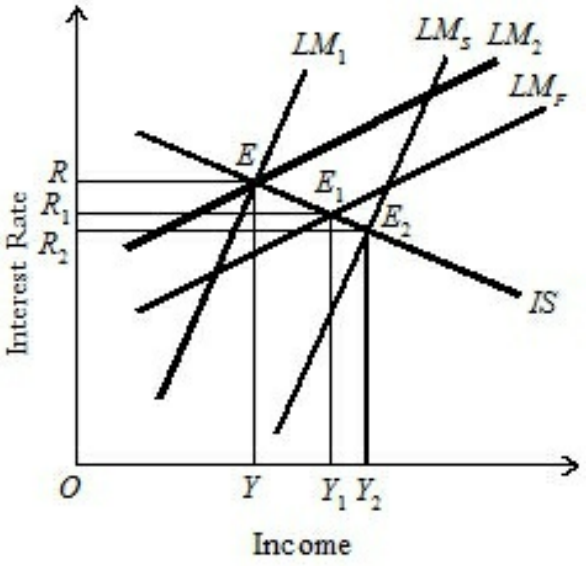


Fig. 2

relative effectiveness of monetary policy depends on the shape of the *LM* curve and the *IS* curve. *Monetary policy is more effective if the LM curve is steeper*. A steeper *LM* curve means that the demand for money is less interest elastic. The less interest elastic is the demand for money, the larger is the fall in interest rate when the money supply is increased. This is because when the demand for money is less elastic to a change in interest rate, an increase in the money supply is more powerful in the bringing about a large fall in interest rate. A large fall in the interest rate leads to a higher increase in investment and in national income. This is depicted in Figure2 where *E* is the original equilibrium position of the economy with *OR* interest rate and *OY* income. When the *steep* LM_1 curve shifts to the right to *LMs*, the new equilibrium is set at E_2. As a result, the interest rate falls from *OR* to OR_2 and income rises from *OY* to OY_2. On the other hand, *the flatter is the LM curve, the less effective is monetary*

policy. A flatter *LM* curve means that the demand for money is more interest elastic. The more interest elastic is the demand for money, the smaller is the fall in interest rate when the money supply is increased. A small fall in the interest rate leads to a smaller increase in investment and income. In Figure 2, E is the original equilibrium position with *OR* interest rate and *OY* income. When the flatter LM_2 curve shifts to the right to LM_F the new equilibrium is established at E_1 which produces OR_1 interest rate and OY_1 income level. In this case, the fall in interest rate to OR_1 is less than OR_2 of the steeper *LMs* curve and the increase in income OY_1 is also less than OY_2 of the steeper curve. This shows that monetary policy is less effective in the case of the flatter *LM* curve and more effective in the case of the steeper curve.

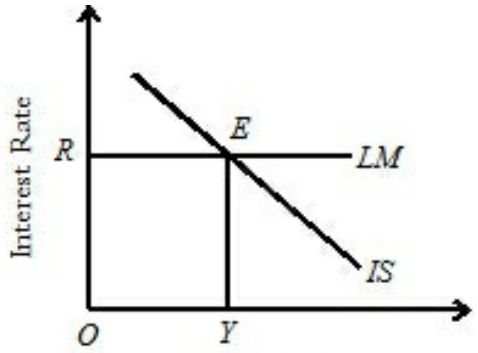


Fig. 3

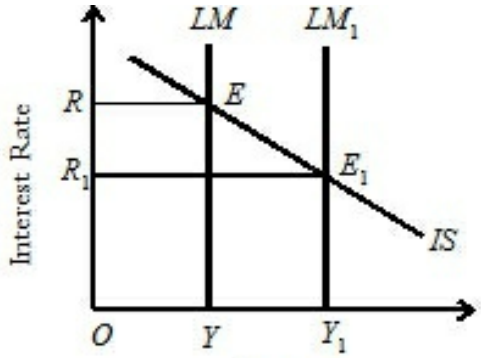


Fig. 4

If *the LM curve is horizontal, monetary policy is completely ineffective* because the demand for money is perfectly interest elastic. This is the case of "liquidity trap" shown in Figure 3, where the increase in the money supply has no effect on the interest rate *OR* and the income level *OY*. On the other hand, if *the LM curve is vertical, monetary policy is highly effective* because the demand for money is perfectly interest inelastic. Figure 4 shows that when the vertical *LM* curve shifts to the right to LM_1 with the increase in the money supply, the interest rate falls from *OR* to OR_1 which has no effect on the demand for money and the entire increase in the money supply has the effect of raising the income level from *OY* to OY_1.

Now take the slope of the *IS* curve. *The flatter is the IS curve, the more effective is the monetary policy*. The flatter *IS* curve means that the investment expenditure is highly interest elastic. When an increase in the money supply lowers the interest rate even slightly, private investment also increases, by a large amount, thereby raising income much. This is depicted in Figure 5 where the original equilibrium is at point *E* with *OR* interest rate and *OY* income level. When the *LM* curve shifts to the right to LM_1 with the increase in money supply, it intersects the flatter curve IS_F at E_2 which produces OR_2 interest rate and OY_2 income. If we compare this equilibrium position E_2 with the E_1 position where the curve ISs is steeper, the interest rate OR_1 and the income level OY_1 are lower than the interest rate and income level of the flatter IS_F curve. This shows that when the money supply is increased, a small fall in the rate of interest leads to a large rise in private in-vestment which raises income more (by YY_2) with the flatter *IS* curve as compared to the steep *IS* curve (by YY_1) thus making monetary policy more effective.

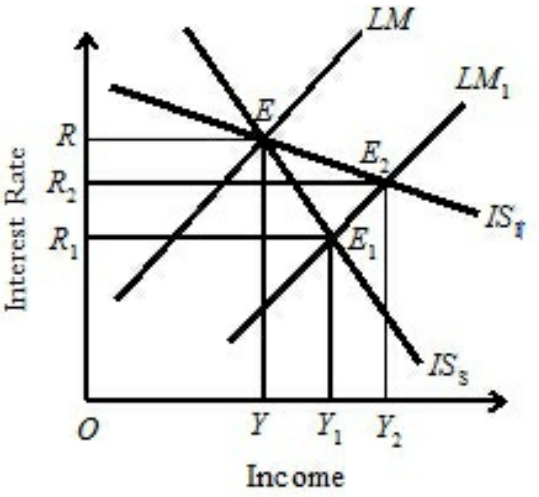


Fig. 5

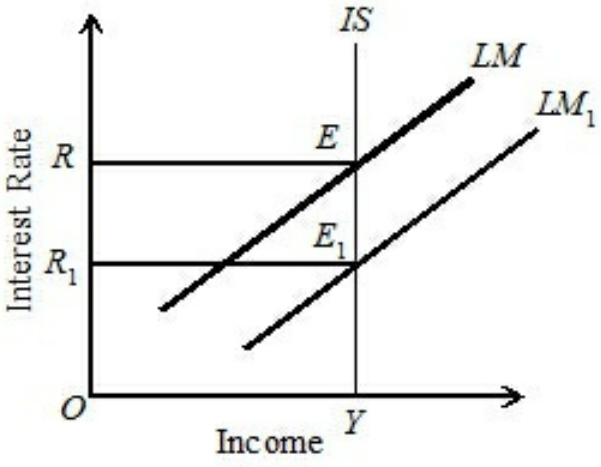


Fig. 6

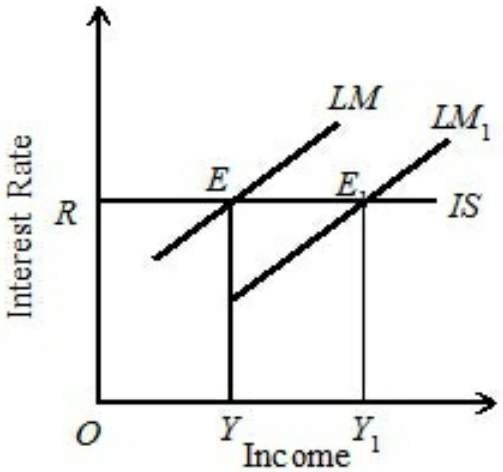


Fig. 7

If the *IS curve is vertical, monetary policy is completely ineffective*

because investment expenditure is completely interest inelastic. With the increase in the money supply, the *LM* curve shifts to the right to LM_1 in Figure 6, the interest rate falls from *OR* to OR_1 but investment being completely interest inelastic, the income remains unchanged at *OY*. On the other hand, if the *IS curve is horizontal, monetary policy is highly effective* because investment expenditure is perfectly interest elastic. Figure 7 shows that with the increase in the money supply, the *LM* curve shifts to LM_1. But even with no change in the interest rate *OR*, there is a large change in income from *OY* to OY_1. This makes monetary policy highly effective.

FISCAL POLICY

The government also influences investment, employment, output and income in the economy through fiscal policy. For an expansionary fiscal policy, the government increases its expenditure or/and reduces taxes. This shifts the *IS* curve to the right. The government follows a contractionary fiscal policy by reducing its expenditure or/and increasing taxes. This shifts the *IS* curve to the left.

Figure 8 illustrates an expansionary fiscal policy with given *IS* and *LM* curves. Suppose the economy is in equilibrium at point *E* with *OR* interest rate and *OY* income. An increase in government spending or a decrease in taxes shifts the *IS* curve upwards to IS_1 which intersects the *LM* curve at E_1. This raises the national income from *OY* to OY_1. The rise in the national income increases the demand for money, given the fixed money supply. This, in turn, raises the interest rate from *OR* to OR_1. The increase in the interest rate tends to reduce private investment expenditure at the same time when the government expenditure is being increased. If the interest rate had not changed with the increase in

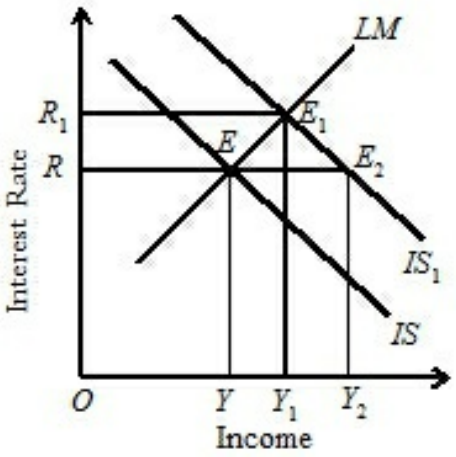


Fig. 8

government expenditure, income would have risen to OY_2 level. But the acutal increase in income has been less by Y_2Y_1 due to the increase in the interest rate to OR_1 which has reduced private investment expenditure. The opposite happens in a contractionary fiscal policy.

The relative effectiveness of fiscal policy depends on the slope of the *LM* curve and the *IS* curve. *Fiscal policy is more effective, the flatter is the LM curve,* and is *less effective when the LM curve is steeper*. When the *IS* curve shifts upwards to IS_1 with the increase in government expenditure, its impact on the national income is more with the flatter LM_F curve than with the steeper LM_S curve. This is shown in Figure 9 where the IS_1 curve intersects the flatter LM_F curve at point E_2 which produces OY_2 income and OR_2 interest rate. On the other hand, it intersects the steeper LM_S curve at E_1 which determines OY_1 income and OR_1 interest rate. In the case of the steeper curve LM_S, the increase in income to OY_1 leads to a large rise in the demand for money which raises the interest rate to a very high level OR_1. The large increase in the interest rate reduces private investment despite increase in government expenditure which ultimately brings a small rise in income OY_1. But in the case of the flatter curve LM_F, the rise in the interest rate to OR_2 is relatively small. Consequently, it reduces private investment to a lesser degree and its net effect on national income is relatively large. Thus the increase in national income with the flatter curve LM_F is more ($YY_2 > YY_1$) as compared with the steeper curve LM_S.

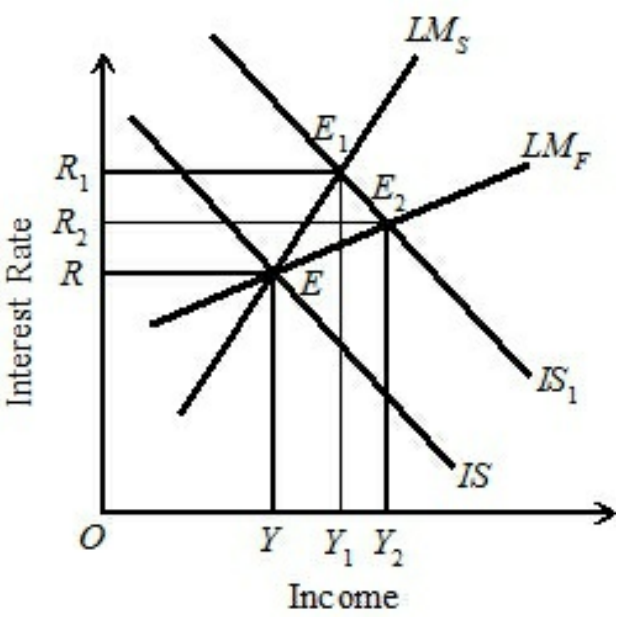


Fig. 9

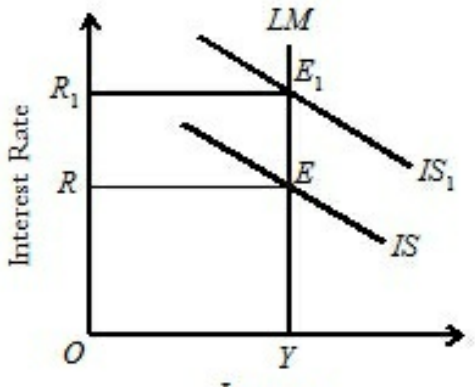


Fig. 10

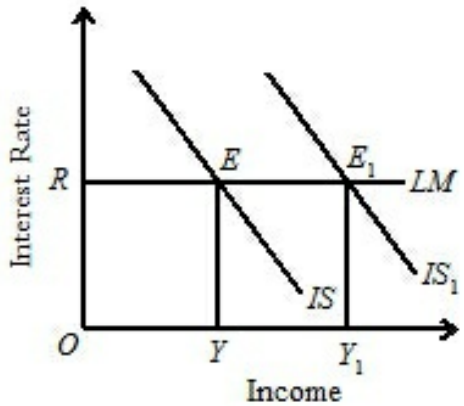


Fig. 11

Fiscal policy is completely ineffective, if the LM curve is vertical. It means that the demand for money is perfectly interest inelastic. This is shown in Figure 10 where the level of income remains unchanged. When the *IS* curve shifts upwards to IS_1, only the interest rate rises from *OR* to OR_1 and increase in government expenditure does not affect national income at all. It remains constant at *OY*. At the other extreme is the *perfectly horizontal LM curve* where *fiscal policy is fully effective*. This situation implies that the demand for money is perfectly interest elastic. This is shown in Figure 11 where the horizontal *LM* curve is intersected by the *IS* curve at *E* which produces *OR* interest rate and *OY* income. When the *IS* curve shifts to the right to IS_1 , income rises by the full multiplier of the increase in government expenditure. It rises to OY_1 but there is no change in interest rate.

Now take the slope of the *IS* curve. *The steeper is the IS curve, the more effective is fiscal policy. The flatter is the IS curve, the less effective is fiscal policy.* These two cases are illustrated in Figure 12 where *E* is the original equilibrium point with *OR* interest rate and *OY* income level. The increase in government expenditure shifts the flatter curve IS_1 to IS_F so that the new equilibrium with *LM* curve at point E_1 produces OR_1 interest rate and OY_1 income level. Similarly, the steeper curve IS_2 is shifted to IS_S with the increase in government expenditure and the new equilibrium with *LM* curve at point E_2 leads to OR_2 interest rate and OY_2 income level. The figure shows that the national income increases more with the shifting of

the steeper *IS* curve than in the case of the flatter *IS* curve. It rises by YY_2 in the case of the steeper curve IS_S and by YY_1 in the case of the flatter curve IS_F. This is because investment expenditure is less interest-elastic, when the *IS* curve is steeper. The increase in the interest rate to OR_2 reduces very little private investment with the result that the rise in income is greater. It is YY_2. On the other hand, the increase in income is smaller in the case of the flatter *IS* curve. It is YY_1. This is because investment expenditure is more interest-elastic. The increase in the interest rate to OR_1 reduces large private investment so that the rise in income is smaller. Thus fiscal policy is more effective, the steeper is the *IS* curve and is less effective in the case of the flatter *IS* curve.

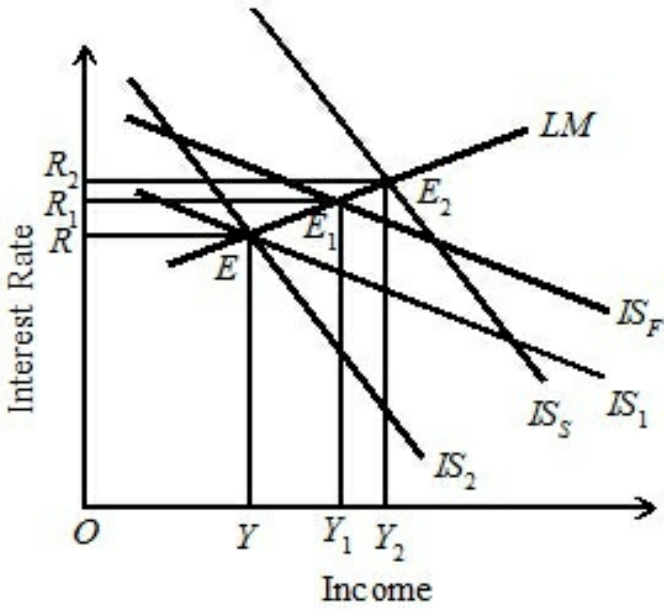


Fig. 12

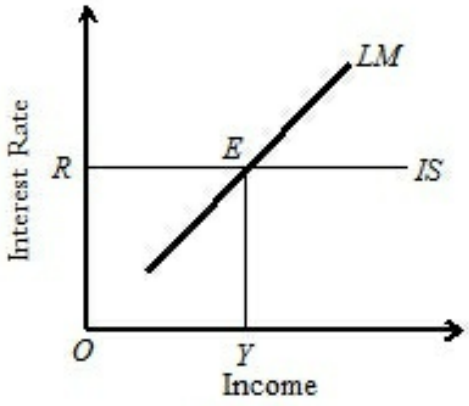


Fig. 13

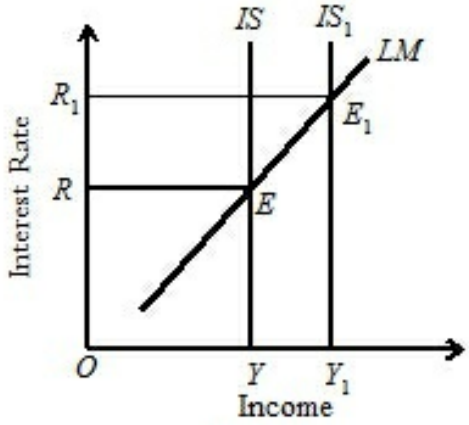


Fig. 14

Fiscal policy is completely ineffective, if the IS curve is horizontal. An horizontal *IS* curve means that investment expenditure is perfectly interest elastic. This is depicted in Figure 13 where *LM* curve intersects the *IS* curve at *E*. An increase in government expenditure has no effect on the

interest rate *OR* and hence on the income level *OY*. Such a situation is not likely to be in practice. On the other extreme is *the vertical IS curve which makes fiscal policy highly effective.* This is because government expenditure is perfectly interest inelastic. An increase in government expenditure shifts the *IS* curve to the right to IS_1, raises the interest rate to OR_1 and income to OY_1 by the full multiplier of the increase in government expenditure, as shown in Figure 14. This makes fiscal policy highly effective.

THE SYNTHESIST VIEW : THREE RANGE ANALYSIS

Economists have explained the effectiveness of monetary and fiscal policies in three ranges in order to reconcile the *extremes* of the Keynesian and monetarist (or classical) views. The *LM* curve slopes upward to the right and has three segments, as shown in Figure 15. Starting from the left it is perfectly elastic, from R_1 to *A*. This segment is known as "*the Keynesian range*", reflecting the "liquidity trap". At the other extreme to the right, the *LM* curve is perfectly inelastic, from *E* to LM_2. This segment of the curve is known as "*the classical range*," because the classicals believed that money is held only for transactions purposes and nothing is held for speculative purposes. In between these two segments of the curve is "*the intermediate range*". The Keynesian range represents the fiscalist or Keynesian view, the classical range the monetarist view, and the intermediate range the synthesist view.

We take expansionary monetary and fiscal policies in order to explain their effectiveness which depend upon the extent to which they affect the level of income and the rate of interest in the Keynesian, the classical and the intermediate ranges. They, in turn, are determined by the responsiveness of the demand for money to changes in the interest rate.

MONETARY POLICY

Monetary policy is explained in Figure 15 where the three-range *LM* curves LM_1 and LM_2 are shown with three *IS* curves. The LM_2 curve

emerges after an increase in the money supply.

The Keynesian Range

First, consider the Keynesian range where the *LM* curve is perfectly elastic. This is the Keynesian liquidity trap situation in which the *LM* curve is horizontal from R_1 to *A*, and the interest rate cannot fall below OR_1. An increase in the money supply shifts the *LM* curve from LM_1 to LM_2. This shift in the curve has no effect on the rate of interest. Consequently, investment is not affected at all so that the level of income remains unchanged at OY_1. This is because at a very low rate of interest such as OR_1 , people prefer to keep money in cash rather than in bonds (or securities) in the hope of converting it into bonds when the interest rate rises. Thus under the Keynesian assumption of the liquidity trap, the horizontal portion of the *LM* curve is not affected by an increase in the money supply. The *IS* curve intersects the *LM* curve in the flat range at A with little effect on the interest rate, investment and income. Monetary policy is, therefore, totally ineffective in the Keynesian range.

The Classical or Monetarist Range

Consider the classical range where *LM* curve is perfectly inelastic. In the classical range, the system is in equilibrium at *D* where the IS_3 curve intersects the LM_1 curve and the interest rate is OR_5 and income level OY_4. Suppose the central bank adopts an expansionary monetary policy whereby it increases the money supply by open market operations. The increase in money supply shifts the LM_1 curve to the right to LM_2 position. As a result, the income level increases from OY_4 to OY_5 and the interest rate falls from OR_5 to OR_4 when the IS_3 curve crosses the LM_2 curve at *E*.

The increase in the income level and fall in the interest rate as a result of the increase in the money supply is based on the classical assumption that money is primarily a medium of exchange. When the central bank buys securities in the market, the security prices are bid up and the rate of interest falls. The wealth holders then find other assets more attractive

than securities. They, therefore, invest the increased cash holdings in new or existing capital investments which, in turn, raise the level of income. But as long as wealth holders possess more money balances than are required for transactions purposes, they will continue to compete for earning assets. Consequently, the interest rate will continue to fall and investment will continue to rise until the excess money balances are absorbed in such transactions. Ultimately, the equilibrium level of income rises by the full amount of the increase in the money supply. Thus the monetary policy is highly effective in the classical range when the economy is at high levels of income and interest rate and utilises the entire increase in the money supply for transactions purposes thereby raising national income by the full increase in the money supply.

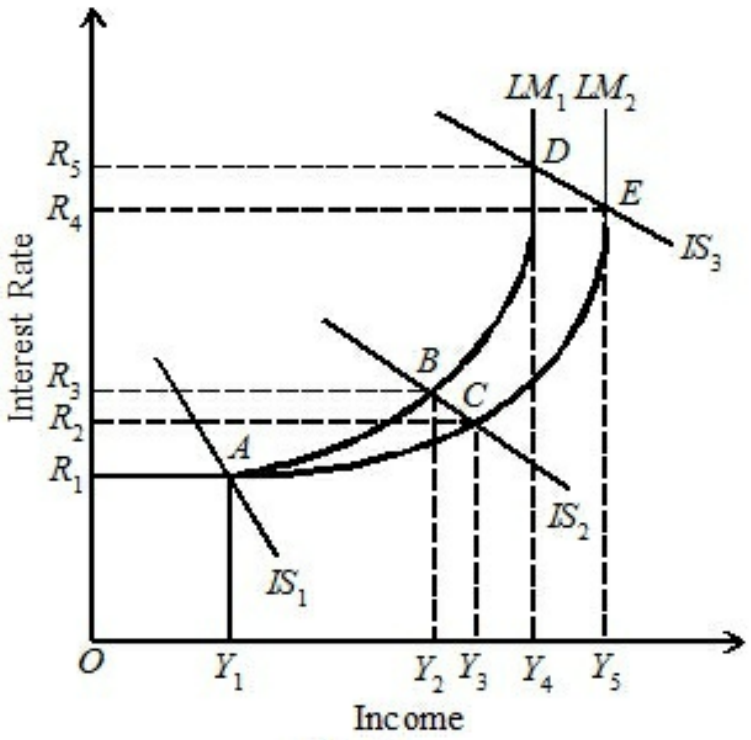


Fig. 15

The Intermediate Range

Now consider the intermediate range when the initial equilibrium is at *B* where the IS_2 curve intersects the LM_1 curve, and the income level is OY_2 and the interest rate is OR_3. The increase in the money supply shifts the LM_1 curve to LM_2 position. As a result, the new equilibrium is established at point *C* where the IS_2 curve crosses the LM_2 curve. It shows that with the increase in the money supply, the rate of interest falls from OR_3 to OR_2 and the income level rises from OY_2 to OY_3. In the intermediate range, the increase in income by Y_2Y_3 is less than that in the classical range, $(Y_2Y_3 < Y_4Y_5)$. This is because in the classical case the entire

increase in the money supply is absorbed for transactions purposes. But in the intermediate case, the increased money supply is partly absorbed for speculative purposes and partly for transactions purposes. That which is held for speculative purposes is not invested by wealth holders and remains with them in the form of idle balances. This has the effect of raising the income level by less than the increase in the money supply. Thus *in the intermediate range monetary policy is less effective than in the classical range.*

FISCAL POLICY

Fiscal policy is explained in Figure 16 in which the three range *LM* curve is taken along with six *IS* curves that arise after increase in government expenditure in the case of the Keynesian, intermediate and classical ranges.

The Keynesian Range

Consider first the Keynesian range when the initial equilibrium is at *A* where the IS_1 curve intersects the *LM* curve. Suppose the government expenditure is increased. This brings about new equilibrium at *B* where the IS_2 curve cuts the *LM* curve. Consequently, the income level rises from OY_1 to OY_2 with the interest rate unchanged at *OR*. The increase in income in the Keynesian case is equal to the full multiplier times the increase in government expenditure. This is because with fixed money supply at low levels of interest rate and income, there is lot of idle money with the wealth holders. This can be used to finance higher transactions without raising the interest rate. When the interest rate does not rise the level of investment remains the same as before and the increase in income is equal to the full multiplier times the increase in government expenditure. *Thus in the Keynesian range, the fiscal policy is very effective.*

The Classical or Monetarist Range

In the classical range, the *LM* curve is perfectly inelastic and the IS_5 curve

intersects it at *E* so that the interest rate is OR_3 and the income level is OY_5. When the government expenditure increases for an expansionary fiscal policy, the IS_5 curve shifts upward to IS_6. As a result, the IS_6 curve crosses the *LM* curve at *F* and the interest rate rises to OR_4 with income remaining unchanged at OY_5. This is because the classical case relates to a fully employed economy where the increase in government expenditure has the effect of raising the interest rate which reduces private investment. Since the increase in government expenditure exactly equals the reduction in the private investment, there is no effect on the level of income which remains constant at OY_5. Thus *fiscal policy is not at all effective in the classical range.*

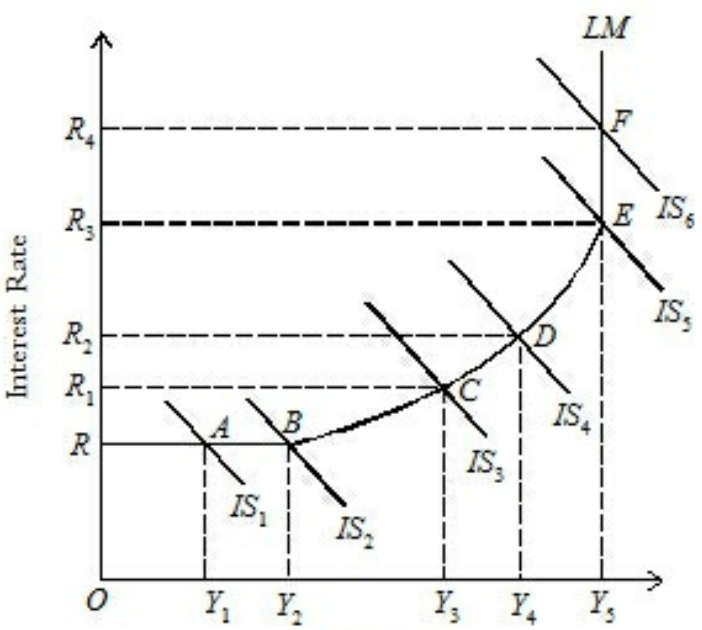


Fig. 16

The Intermediate Range

In the intermediate range, the initial equilibrium is at *C* where the IS_3 curve intersects the *LM* curve. Here OR_1 is the interest rate with OY_3 income level. With the increase in the government expenditure, the IS_3 curve shifts upward to the right to IS_4 and the new equilibrium between IS_4 and *LM* curves is established at point *D*. As a result, the increase in government expenditure raises the income level from OY_3 to OY_4 and the interest rate from OR_1 to OR_2. The increase in both the income level and the interest rate in the intermediate range is due to two reasons. *First*, the increase in income resulting from a rise in government expenditure occurs because additional money balances are available for transactions purposes. *Second*, given a fixed money supply, a part of available transactions are

held as idle balances by wealth holders which raise the interest rate. As a result of the rise in the interest rate, investment falls and the *fiscal policy is not so effective as in the Keynesian range*. In general, fiscal policy will be more effective the closer equilibrium is to the Keynesian range and less effective the closer equilibrium is to the classical range.

Effects of Elasticities of *IS* Curve on Monetary and Fiscal Policies

The elasticities of the *IS* curve affect monetary and fiscal policies in a slightly different way. This is explained in terms of Figure 17. In the *Keynesian range, monetary policy is ineffective whether the IS curve is elastic or inelastic*. On the other hand, *fiscal policy is only effective when the IS curve is elastic or inelastic*. The elastic curve $IS_{\mathbf{F}}$ shifts to $IS_{\mathbf{F1}}$ and income rises from $OY_{\mathbf{1}}$ to $OY_{\mathbf{2}}$ in Figure 17. The same result follows in the case of the shifting of an inelastic *IS* curve (not shown in figure).

In the *classical range, fiscal policy is ineffective whether the IS curve is elastic* (*IS*F2) *or inelastic* ($IS_{\mathbf{S2}}$). *But monetary policy is effective under both the elastic and inelastic curves*. Income rises from $OY_{\mathbf{3}}$ to $OY_{\mathbf{6}}$, as shown in Figure 17.

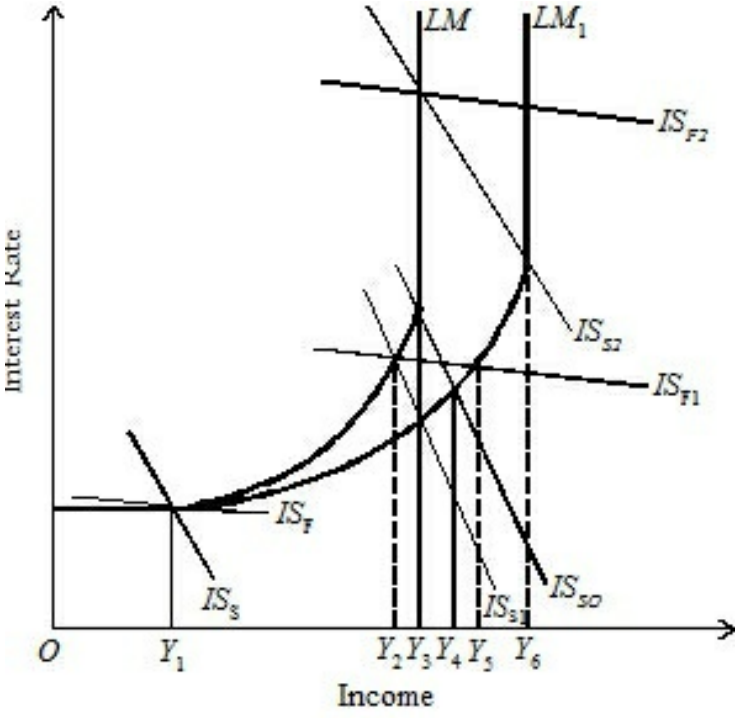


Fig. 17

In the intermediate range, *monetary policy* is *less* effective when the $IS_{\mathbf{S1}}$ curve is *inelastic* because the rise in income in this case is $Y_{\mathbf{2}}Y_{\mathbf{3}}$ whereas in the case of the *elastic* curve $IS_{\mathbf{F1}}$, it is *more effective,* the rise in income being $Y_{\mathbf{2}}\ Y_{\mathbf{5}}(>Y_{\mathbf{2}}Y_{\mathbf{3}})$. But *fiscal* policy is *more effective*, whether the *IS* curve is *elastic* or *inelastic*. The shifting of the inelastic curve $IS_{\mathbf{S1}}$ to IS_{SO} shows

the increase in income from OY_3 to OY_4.

Conclusion

The relative effectiveness of monetary and fiscal policy depends upon the shape of the *IS* and *LM* curves and the economy's initial position. If the economy is in the Keynesian range, monetary policy is ineffective and fiscal policy is highly effective. On the other hand, in the classical range, monetary policy is effective and fiscal policy is ineffective. But in the intermediate range both monetary and fiscal policies are effective. This case bridges the gap between the Keynesian and classical views. In this range, the elasticities of the *IS* and *LM* curves are neither highly interest elastic nor highly interest inelastic. In fact, in the intermediate range, the effectiveness of monetary and fiscal policies depends largely on the elasticities of the *IS* curve. If the *IS* curve is inelastic, fiscal policy is more effective than monetary policy. On the other hand, if the *IS* curve is elastic, monetary policy is more effective than fiscal policy. Thus for a complete effectiveness of both monetary and fiscal policies the best course is to have a monetary-fiscal mix.

MONETARY-FISCAL MIX

Consider a situation where an expansionary mix of monetary-fiscal policies is adopted to achieve full employment in the economy. This is illustrated in Figure 18 where the economy is in the initial situation at *A* on the basis of the interaction of IS_1 and LM_1 curves. This situation depicts OR_2 interest rate and OY_1 income level. Now an expansionary fiscal policy is adopted in the form of increase in government expenditure or decrease in taxes. This shifts the curve IS_1 to IS_2. This will have the effect of raising the interest rate further to OR_3 if an expansionary

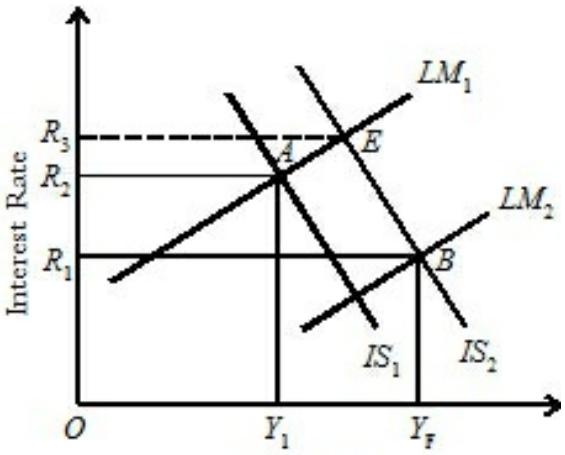


Fig. 18

monetary policy is not adopted simultaneously. So in order to reduce the interest rate and encourage investment for achieving full employment, the monetary authority increases the money supply through open market purchase of securities. This tends to shift the curve LM_1 to the right in the position of LM_2 curve. Now fiscal policy has led to the new IS_2 curve and monetary policy to the *LM2* curve. Both the curves intersect at *B* whereby the interest rate is lowered to OR_1 and the level of income rises to the full employment level OY_F.

Let us take another situation when the economy is at the full employment level of income OY_F where the *IS* curve intersects the *LM* curve at point *E* in Figure 19. But due to some reasons, the economy's growth rate has slowed down. In order to overcome this, more investment is required to be made in the economy. For this, the monetary authority increases the money supply which leads to the shifting of the curve *LM* to the right to LM_1. The LM_1 curve intersects the *IS* curve at point E_1 which lowers the interest rate to OR_1 and raises the income level to OY_1. But the rise in national income being higher than the full employment income level, this policy is inflationary. Therefore, the economy requires a change in the monetary-fiscal policy mix.

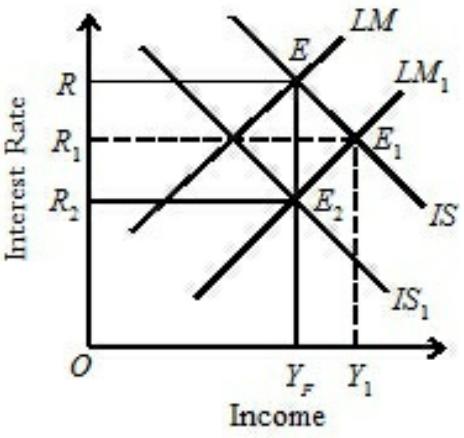


Fig. 19

For this, the expansionary monetary policy should be combined with a restrictive fiscal policy. Accordingly, the government reduces its investment expenditure or/and increases taxes so that the *IS* curve shifts to the left to IS_1. Now the IS_1 curve intersects the LM_1 curve at point E_2 so that the new equilibrium is established at a lower interest rate OR_2 and income level OY_F which is the full employment income level. This level can be maintained by the present monetary-fiscal policy mix because the lower interest rate would keep large investment spending in the economy and reduced government expenditure or high taxes would control

inflation.

EXERCISES

1. Examine the effectiveness of monetary and fiscal policies in terms of the *IS* and *LM* functions.
2. Examine the extent of effectiveness of monetary and fiscal policies in counteracting recessionary forces.
3. How would you reconcile the extremes of the Keynesian and classical views on the effectiveness of monetary and fiscal policies ?
4. Explain the monetary-fiscal mix to achieve and maintain full employment.

PART-VIII

MODERN MACROECONOMICS

CHAPTER 54

THE RATIONAL EXPECTATIONS HYPOTHESIS

INTRODUCTION

In the 1930s when Keynes wrote his *General Theory*, unemployment was the major problem in the world. During the Second World War, inflation emerged as the main economic problem. In the postwar years till the late 1960s, unemployment again became a major economic issue. From the late 1960s to 1970s, a new phenomenon appeared in the form of both high unemployment and inflation, known as *stagflation*. This phenomenon of stagflation posed a serious challenge to economists and policy makers because the Keynesian theory was silent about it. Out of this crisis emerged a new macroeconomic theory which is called the Rational Expectations Hypothesis (Ratex).

ADAPTIVE EXPECTATIONS

Before we discuss the Ratex hypothesis, it is essential to understand the meaning of adaptive expectations used in macroeconomics before the Ratex theory was developed.

Expectations are forecasts or predictions by an economic agent regarding the uncertain economic variables which are relevant to his decision. They are based on post trends as well as current information and experience. The main body of economic theory is based on the assumption of rational behaviour of economic agents (i.e., consumers, producers, etc.) in forming their expectations. But till recently economists have not been able to incorporate the role of expectations in measuring human behaviour.

Keynes discussed the *importance* of expectations but he was silent as to how they are formed.

In recent years, economists have mostly used the adaptive expectations hypothesis in model building. The pioneering work was done by Cagan[1] in 1956 and Nerlove[2] in 1957. According to the adaptive expectations hypothesis, economic agents expect the future to be essentially a continuation of the past. They expect the future values of economic variables like prices, incomes, etc. to be an *average* of past values and to change very slowly. The economic agents make the expected values of these variables equal to a *weighted* average of their present and past values. They revise their expectations in accordance with the last forecasting error. Errors resulting from past behaviour represent an important source of information for forming expectations. But such expectations are based on the assumption that the economic agents expect them to change very little. This often leads to absurd forecasts when there is change in economic policy.

For instance, according to the adaptive expectations hypothesis, economic agents form expectations of future inflation rates from a weighted average of experienced average past inflation rates and they periodically revise those expectations if actual inflation turns out to be different than expected. This implies irrational behaviour on the part of economic agents. Friedman's analysis of the long-run Phillips curve is based on the adaptive expectations hypothesis. The assumption implicit in Friedman's *acceleration* hypothesis that price expectations are based mainly on the basis of the experience of past inflation is unrealistic. When economic agents base their price expectations on this assumption, they are irrational. If they think like this in a period of rising prices, they will find that they were wrong. This is because expectations are formed from direct forecasts of the future as from mere projections of the past. People base their expectations as much on current information about a variety of factors as on past price changes. Thus rational people will use all available information to forecast future inflation more accurately.

RATIONAL EXPECTATIONS

The idea of rational expectations was first put forth by Johy Muth[3] in 1961 who borrowed the concept from engineering literature. His model dealt mainly with modelling price movements in markets. By assuming that economic agents optimise and use information efficiently when forming expectations, he was able to construct a theory of expectations in which consumers' and producers' responses to *expected* price changes depended on their responses to *actual* price changes. Muth pointed out that certain *expectations are rational in the sense that expectations and events differ only by a random forecast error.*

---

1. P. Cagan, "The Monetary Dynamics of Hyperinflation," in M. Friedman (ed.), *Studies in the Quantity Theory of Money*, 1956.

2. M. Nerlove, "Adaptive Expectations and Cobweb Phenomena," *O.J.E.*, May 1958.

3. John F. Muth, "Rational Expectations and the Theory of Price Movements" *Econometrica*, July 1961.

Muth's notion of rational expectations related to microeconomics. It did not convince many economists and lay dormant for ten years. It was in early 1970s that Robert Lucas, Thomas Sargent and Neil Wallace applied the idea to problems of macroeconomic policy.

Basic Propositions of the Rational Expectations Hypothesis

The Ratex hypothesis holds that economic agents form expectations of the future values of economic variables like prices, incomes, etc. by using all the economic information available to them. This information includes the relationships governing economic variables, particularly monetary and fiscal policies of the government. Thus the rational expectationists assume that economic agents have full and accurate information about future economic events. According to Muth, information should be considered like any other available resource which is scarce. Further, rational economic agents should use their knowledge of the structure of the economic system in forming their expectations. Thus the Ratex hypothesis "presumes that individual economic agents use all available and relevant information in forming expectations and that they process this information in an intelligent fashion. It is important to recognise that this does not imply that consumers or firms have "perfect foresight" or that their

expectations are always "correct". What it does suggest is that agents reflect upon past errors and, if necessary, revise their expectational behaviour so as to eliminate regularities in these errors. Indeed the hypothesis suggests that agents *succeed* in eliminating regularities involving expectational errors, so that the errors will on the average be unrelated to available information."

The Ratex hypothesis has been applied to economic (monetary, fiscal and income) policies. The rational expectationists have shown the short-run ineffectiveness of stabilisation policies. According to them, no one knows much about what happens to the economy when economic (monetary or fiscal) policy is changed. Specifically, it means that macroeconomic policies designed to control recession by cutting taxes, increasing government spending, increasing the money supply or the budget deficit may be curbed. They argue that the public has learnt from the past experience that the government will follow such a policy. Therefore, the government cannot fool the people by adopting its effects and mere signs of such a policy in the economy create expectations of countercyclical action on the part of the public. Thus, according to the Ratex hypothesis, people form expectations about government monetary and fiscal policies and then refer to them in making economic decisions. As a result, by the time signs of government policies appear, the public has already acted upon them, thereby offsetting their effects. In other words, the Ratex hypothesis holds that the only policy moves that cause changes in people's economic behaviour are those that are not expected, the surprise moves by the government. Once the public acquires knowledge about a policy and expects it, it cannot change people's economic behaviour. We discuss some of the policy changes in the light of the Ratex hypothesis below.

Rational Expectations and the Phillips Curve

In the Friedman-Phelps acceleration hypothesis of the Phillips curve, there is a short-run trade-off between unemployment and inflation but no long-run trade-off exists. The reason is that inflationary expectations are based on past behaviour of inflation which cannot be predicted accurately. Therefore, there is always an observed error so that the expected rate of inflation always lags behind the actual rate. But the expected rate of

inflation is revised in accordance with the first period's experience of inflation by adding on some proportion of the observed error in the previous period so that the expected rate of inflation adjusts toward the actual rate.

Economists belonging to the rational expections school have denied the possibility of any trade-off between inflation and unemployment even during the long run. According to them, the assumption implicit in Friedman's version that price expectations are formed mainly on the basis of the experience of past inflation is unrealistic. When people base their price expectations on this assumption, they are irrational. If they think like this during a period of rising prices, they will find that they were wrong. But rational people will not commit this mistake. Rather, they will use all available information to forecast future inflation more accurately.

The rational expectations idea is explained diagrammatically in Figure 1 in relation to the Phillips curve. Suppose the unemployment rate is 3 per cent in the economy and the inflation rate is 2 per cent. We start at point A on the SPC_1 curve. In order to reduce unemployment, the government increases the rate of money supply so as to stimulate the economy. Prices start rising. According to the Ratex hypothesis, firms have better information about prices in their own industry than about the general level of prices. They mistakenly think that the increase in prices is due to the increase in the demand for their products. As a result, they employ more workers in order to increase output. In this way, they reduce unemployment. The workers also mistake the rise in prices as related to their own industry. But wages rise as the demand for labour increases and workers think that the increase in money wages is an increase in real wages. Thus the economy moves

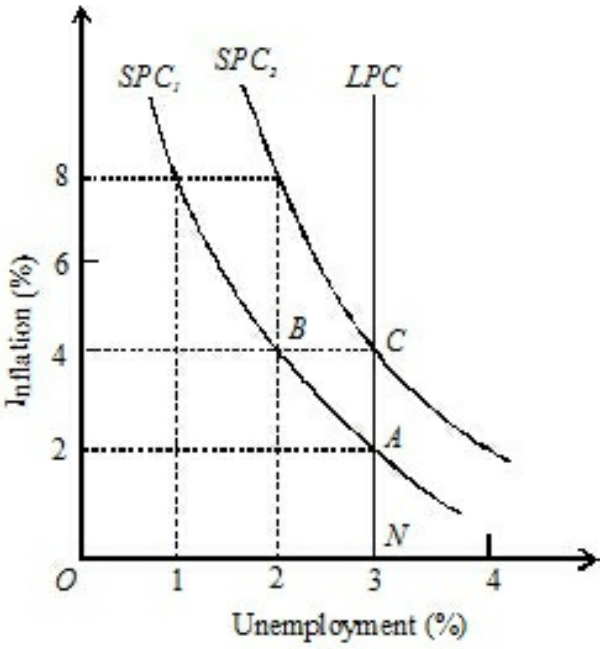


Fig. 1

upward on the short-run Phillips curve SPC_1 from point *A* to *B*. But soon workers and firms find that the increase in prices and wages is prevalent in most industries. Firms find that their costs have increased. Workers realise that their real wages have fallen due to the rise in the inflation rate to 4 per cent and they press for increase in wages. Thus the economy finds itself at the higher inflation rate due to government's monetary policy. As a result, it moves from point *B* to point *C* on the SPC_2 curve where the unemployment rate is 3 per cent which is the same before the government adopted an expansionary monetary policy

When the government again tries to reduce unemployment by again increasing the money supply, it cannot fool workers and firms who will now watch the movements of prices and costs in the economy. If firms expect higher costs with higher prices for their products, they are not likely to increase their production, as happened in the case of the SPC_1 curve. So far as workers are concerned, labour unions will demand higher wages to keep pace with prices moving up in the economy. When the government continues an expansionary monetary (or fiscal) policy, firms and workers get accustomed to it. They build their experience into their expectations. So when the government again adopts such a policy, firms raise prices of their products to nullify the expected inflation so that there is no effect on production and employment. Similarly, workers demand higher wages in expectation of inflation and firms do not offer more jobs. In other words, firms and workers build expectations into their price policies and wage agreements so that there is no possibility for the actual rate of unemployment to differ from the natural rate, *N*, even during the short run.

Its Policy Implications

The Ratex hypothesis assumes that people have all the relevant information of the economic variables. Any discrepancy between the actual rate of inflation and the expected rate is only in the nature of a random crror. When people act rationally, they know that past increases in prices and the rate of change in prices have invariably been accompanied by equal proportional changes in the quantity of money. When people act on this knowledge, it leads to the conclusion that there is no trade-off

between inflation and unemployment even in the short-run. It implies that monetary (or fiscal) policy is unable to change the difference between the actual and natural rate of unemployment. This means that the economy can only be to the left or right of point N of the long-run Phillips curve *LPC* (in Figure 1) in a random manner. Thus the implication is that stabilisation policy is ineffective and should be abandoned.

Stabilisation Policy and Ratex Hypothesis

According to the Ratex hypothesis, monetary and fiscal (stabilisation) policies are ineffective even in the short-run because it is not possible to anticipate accurately how expectations are formed during the short-run. This is called "policy impotence." The Ratex hypothesis is based on the assumption that consumers and firms have accurate information about future economic events. Their expectations are rational because they take into account all available information, especially about expected government actions. If the government is following any consistent monetary or fiscal policy, people know about it and adjust their plans accordingly. So when the government adopts the expected policy measure, it will not be effective because it has been anticipated by the people who have already adjusted their plans. This means that government policy is *ineffective*. Another important assumption is that all markets are fully competitive and prices and wages are completely flexible.

Let us first take *fiscal policy*. The Keynesians advocate an "activist" fiscal policy to reduce unemployment. But, according to the Ratex hypothesis, a tax cut and/or increase in government spending will reduce unemployment only if its short-run effects on the economy are *unexpected* (or unanticipated) by people. In other words, an expansionary fiscal policy may have short-term effects on reducing unemployment provided people do not anticipate that prices will rise. But when the government persists will such a policy, people expect the rate of inflation to rise. So the workers will press for higher wages in anticipation of more inflation in the future and firms will raise the prices of their products in anticipation of the rise in future costs. As a result, fiscal policy will become ineffective in the short-run. It may cause more unemployment and inflation in the long-run when the government tries to control inflation.

Similarly, if the government adopts an expansionary *monetary policy* by increasing the money supply to reduce unemployment, it is also ineffective in the short-run. Such a policy may reduce unemployment, in the short-run provided its effects on the economy are unanticipated. But when the government persists with such an expansionary monetary policy, people expect the inflation rate to rise. Firms raise the prices of their products to overcome the anticipated inflation so that there is no effect on production. Similarly, workers press for higher wages in anticipation of inflation and firms do not employ more workers. So there is no effect on employment.

Thus the Ratex hypothesis suggests that expansionary fiscal and monetary policies will have a temporary effect on unemployment and if continued may cause more inflation and unemployment. For such policies to be successful, they must be unanticipated by the people. Once people anticipate these policies and make adjustments towards them, the economy reverts back to the natural rate of unemployment. Thus for expansionary fiscal and monetary policies to have an impact on unemployment in the short-run, the government must be able to fool the people. But it is unlikely to happen all the time. If the government continues to persist with such policies, they become ineffective because people cannot be fooled for long and they anticipate their effects on production and unemployment. Thus fiscal-monetary policies become ineffective in the short-run. According to the advocates of the Ratex hypothesis, inflation can be controlled without causing widespread unemployment, if the government announces fiscal and monetary measures and convinces the people about it and do not take them be surprise.

Its Criticisms

The Ratex hypothesis has been criticised by economists on the following grounds.

1. Unrealistic Assumption. The assumption of rational expectations is unrealistic. The critics argue that large firms may be able to forecast accurately, but a small firm or the average worker will not be able to do so.

2. Costly Information. It costs much to collect, distill and disseminate information. So the market for information is not perfect. Therefore, the majority of economic agents cannot act on the basis of rational expectations.

3. Different Informations. The critics also point out that the information available to the government differs from that available to firms and workers. Consequently, expectations of the latter about the expected rate of inflation need not necessarily be diverse from the actual rate only by the random error. But the government can accurately forecast about the difference between the expected inflation rate and actual rate on the basis of information available with it. Even if both individuals and government have equal access to information, there is no guarantee that their expectations will be rational.

4. Prices and Wages not Flexible. Critics point out that prices and wages are not flexible. Economists like Philips, Taylor and Fischer have shown that if wages and prices are rigid, monetary or fiscal policy becomes *effective* in the short-run. The rigidity of wage rates implies that they adjust to market forces relatively slowly because wage contacts are binding for two or three years at a time. Similarly, the expected price level at the beginning of the period is expected to hold till the end of the period. Thus even if expectations are rational, monetary or fiscal policy can influence production and unemployment in the short-run.

5. Expectations Adaptive. Gordon rejects the logic of the Ratex hypothesis entirely. He assigns two reasons for this: *first*, individuals do not know enough about the structure of the economy to estimate the market clearing price level and stick with adaptive expectations; and *second*, if individuals gradually learn about the structure of economic system by a least-squares learning method, rational expectations closely approximate to adaptive expectations.

6. Government not Impotent. It is generally said that according to the Ratex hypothesis, the government is impotent in the economic sphere. But the Ratex economists do not claim this. Rather, they believe that the government has a tremendous influence on economic policies.

SELECTED READINGS

R.E. Lucas and T.J. Sargent (eds.), *Rational Expectations and Economic Practice*, 1981.

D.K.H. Begg, *The Rational Expections Revolution in Macroeconomics*, 1982.

EXERCISES

1. State the basic propositions of the rational expectations hypothesis.
2. What is the rational expectationists' position on inflation-unemployment trade-off ? What is the implication of this position for stabilisation policy ?
3. Write a note on adaptive expectations.

CHAPTER 55

SUPPLY-SIDE ECONOMICS

INTRODUCTION

Supply-side economics is a relatively new term which came into use in the mid-1970s as a result of the failure of Keynesian demand-side policies in the US economy which led to stagflation. The term is new but its basic principles are to be found in the works of the classical economists. According to J.B. Say, supply creates its own demand. The very act of supplying goods implies a demand for them. If there is an imbalance between demand and supply, it is corrected automatically by changes in prices and wages and the economy always tends toward full employment. The main emphasis of the classical economists was on economic growth for which they advocated non-interference with the market mechanism. It was the "invisible hand" which led to the maximisation of national wealth. They believed that entrepreneurs, investors and producers were the prime movers on which the economy depended. It was the increase in the supplies of capital and labour and increase in their productivities that determined growth. Of course, free trade and capital movements internationally were instrumental in a faster growth rate of the economy.

MAIN FEATURES OF SUPPLY-SIDE ECONOMICS

Modern supply-side economics lays emphasis on providing all types of economic incentives to raise aggregate supply in the economy. According to Bethell, "The essential argument of supply-side theory is that adding to supply unlike adding to demand is not a zero-sum task. In order to make something, ... a producer does not need to be given any money. Instead, he

has to be given an incentive."[1] Incentives to producers are essential to invest, produce and employ. Similar incentives are to be given to individuals to work and save more. The government plays a limited role in liberalising markets, reducing taxes and freeing the labour market. The main objectives of supply-side policies are to keep inflation at a low level, achieve and maintain full employment and attain faster economic growth. Supply-side economists suggest the following policy measures in order to achieve these objectives.

Tax-induced Change in Aggregate Supply

Supply-siders regard tax cuts as an effective means of raising the growth rate of the economy. To assess the likely effects of tax reductions, they distinguish between income and substitution effects of a cut in the marginal rate of income tax. The substitution effect of a wage cut induces people to work more and have less leisure, and the income effect causes people to work less and enjoy more leisure. It is only when the substitution effect of a tax cut is larger than the income effect that there will be an incentive to work more, thereby leading to reduction in unemployment.

A reduction in personal tax rates increases the incentive of people to work and save more. High savings reduce short-term interest rates and lead to increased investment and thus to an increase in the economy's capital stock. Reduction in marginal tax rates by improving the work effort of the people also increases their productive capacity and the level of output and employment in the economy.

Thus supply-side tax cuts by raising work, effort, saving and investment, increase the supplies of labour and capital and shift the aggregate supply curve to the right. The effect of a. supply-side tax cut is illustrated in Fig. 1 where *AS* is the aggregate supply curve and *AD* is the given demand curve. Real output or *GDP* is measured along the horizontal axis and the price level on the vertical axis. *AS* and *AD* curves intersect at point *T* and determine *OP* price and *OQ* real output of the economy. Suppose there is a tax cut both on persons and firms. This increases work effort and saving on the part of workers and investment by firms. As a result, supplies of labour and capital increase which shift the aggregate supply curve AS to

the right as AS_1. Now the AS_1 curve cuts the *AD* curve at point *C.As* a result, the price level falls to OP_1 and the real output increases to QQ_t as a result of a tax cut.

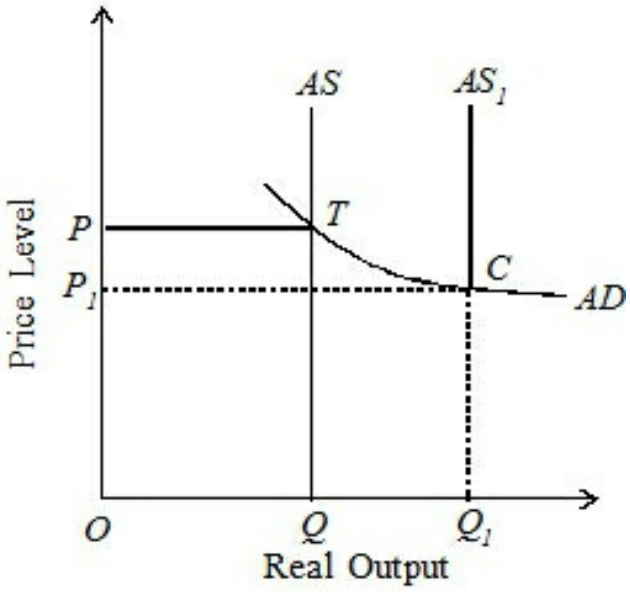


Fig. 1

Similarly, reduction in corporate tax rates, by giving incentives to the corporate sector in the form of increasing tax credit for larger investment and providing higher depreciation allowance, encourage investment. Higher investment leads to the production of more goods and services per unit of labour and capital.

Supply-siders also advocate an additional tax relief for firms employing researchers because R&D help in increasing productivity. They also favour reduced estate taxes for small farmers which will induce them to spend more on inputs so as to increase production.

Further, tax cuts reduce diversions to “shelter” (protected) industries and minimise or eliminate the need for accountants, investment consultants and tax-lawyers. Moreover, tax reductions reduce ‘underground’ (black market) activity where exchange is not recorded in the books and no taxes are paid.

Increasing Growth Rate

According to supply-side economists, tax cuts increase the disposable income of the people who raise additional demand for goods and services. On the other hand, the faster growth in productivity leads to the production of additional goods and services to match the additional demand. This leads to balanced growth in the economy without shortages. When the economy is moving towards balanced growth, the rate of inflation is low. This, in turn, leads to an increase in the real disposable

income of the people which raises consumption, output and employment. Low inflation leads to increase in net exports which strengthens the value of national currency in relation to foreign currencies. The increase in productivity increases the production of more goods for export, thereby further strengthening the country's currency.

Thus supply-side economists advocate reduction in tax rates in order to increase the incentives to work, save and invest and to get more tax revenue by the government. Increase in investment leads to an increase in the economy's capital stock, to increase in productivity, to larger output, low inflation, high level of employment and high growth rate of the economy. These policy prescriptions shift the aggregate supply curve of the economy to the right This is illustrated in Fig. 2 where *AS* is the aggregate supply curve and *AD* is the given aggregate demand curve. They intersect at point *E* which is the initial equilibrium point of the economy with *OP* price level and *OQ* real output. Suppose the supply-side policies increase the total supply of factors like labour and capital due to tax policies, incentives, etc. They increase real output and shift the *AS* curve to the right as AS_1. The new equilibrium is at E_1 where the AS_1 curve cuts the *AD* curve. Now real output increases to OQ_1 and the price level falls to OP_1, thereby increasing the growth rate of the economy.

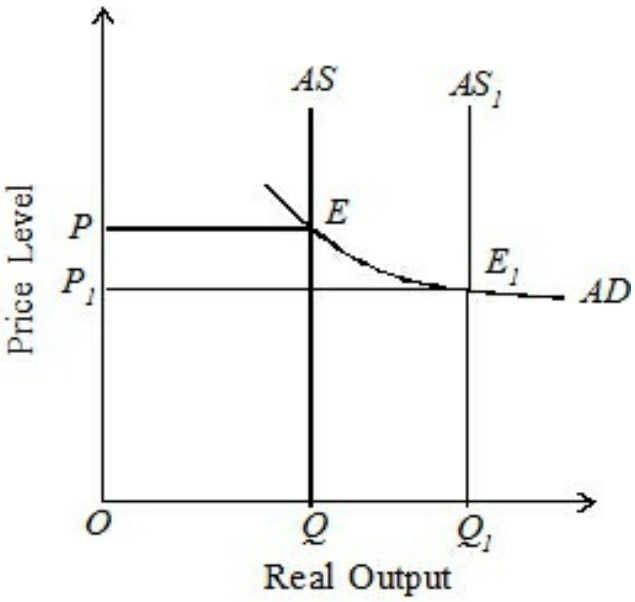


Fig. 2

POLICY PRESCRIPTIONS OF SUPPLY-SIDE ECONOMICS

Following are the policy prescriptions of supply-side economics:

1. The Laffer Curve : Tax Rate Vs Tax Revenue. The most popular

aspect of supply-side economics is the Laffer curve named after its originator Prof. Arthur Laffer. The Laffer curve depicts the relation between tax rate and tax revenue. It is based on the assumption that a cut in the marginal rate of tax will increase the incentives to work, save and invest This tax cut, in turn, will increase the tax revenue. The Laffer curve shows two extremes of tax rates : A 0% tax rate and a 100% tax rate.

Both yield no tax revenue. If the tax rate is 0%, no revenue will be raised. If the tax rate is 100%, people will have no incentive to work, save and invest at all because the whole income will go to the government. Thus the tax revenue will again be zero. As the tax rate increases from 0% to 100%, tax revenue correspondingly rises from zero to some maximum level and then starts declining to zero. Thus the optimum tax rate is somewhere between the two extremes.

Figure 3 shows the Laffer curve where the tax rate (0%) is taken on the horizontal axis and the tax revenue on the vertical axis. As the tax rate is raised above zero, the tax revenue starts increasing. The Laffer curve is upward sloping. At the relatively low tax rate, it is upward sloping. At the relatively low tax rate T_1, the tax revenue is R_1. As the tax rate rises to T, the tax revenue continues toincrease and the curve reaches the peak, P where the tax revenue R is the maximum. Thereafter, further rise in the tax rate will reduce revenue to the government. Thus T is the optimum rate of tax. According to Laffer, “Except for the optimum rate, there are always two tax rates that yield the same revenue.” In the figure, the revenue R_1 at the high tax rate T_2 is the same as the revenue collected at the low tax rate T_1 . If the government wishes to maximise tax revenue, it will choose the optimum tax rate T.

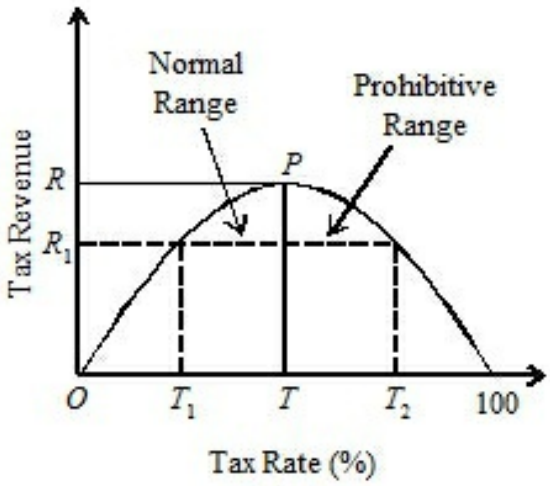


Fig. 3

An important feature of the Laffer curve is that it has a *normal* range and a *prohibitive* range. The normal range is to the left of the optimum tax rate T

and the prohibitive range is to its right. In the normal range, increases in the tax rate bring more revenue to the government. But in the prohibitive range, when the tax rate becomes high, it reduces the incentives to work, save and invest. Consequently, the fall in output more than offsets the rise in tax rate. When the tax rate reaches 100%, the revenue falls to zero because no one will bother to work. Thus high tax rate stifles economic growth and results in high unemployment. Therefore, a reduction in the tax rate will actually increase revenue by encouraging the incentives to work, save and invest. People not only produce and earn more but also switch money out of low-yielding "tax shelters" and untaxed "underground" economy into more productive and socially desirable investment. The result would be higher employment and economic growth leading to high tax revenue.[2]

2. Reduction in Government Spending. To achieve full employment, low inflation and high growth rate of the economy, the supply-side economists emphasise reduction in government expenditure accompanied by tax cuts. They are against monetization of budget deficit[3] which the Keynesians advocate. But the reduction in government expenditure should be more than or equal to tax cuts so that savings increase to finance larger investments. This will increase employment, income and growth rate of the economy.

3. Monetary Policy. Another plank of supply-side policy is to have restrained monetary expansion in order to keep the inflation rate low.

4. Increased Depreciation. To encourage more investment, supply-siders suggest increased investment allowance and/or higher depreciation on buildings, machines vehicles, and other capital goods.

5. Reduction in Welfare Benefits. To reduce unemployment, supply-side economists emphasise reduction in welfare benefits, especially unemployment allowance. This will encourage workers to accept jobs at lower wages, thereby reducing unemployment in the economy.

6. Reducing Trade Union Power. Supply-siders also advocate reduction in the power of trade unions through legislation which will make the labour market more competitive. Trade unions raise wages above the

competitive level which the employers cannot afford. Thus they destroy jobs and increase unemployment. When the government restricts union power, unemployment and cost-inflation are reduced.

7. Deregulation and Privatisation. Deregulation and privatisation are important supply-side policies. They are used to encourage more competition within the economy. Removal of public sector monopolies, and sale of public sector enterprises and transfer of public utilities in private hands lead to increase in productive efficiency, wider consumer choice and lower prices.

8. Free Trade and Capital Movements. Free trade and free capital movements among countries is another policy measure of supply-siders. The removal of exchange controls and free inflow and outflow of both short-term and long-term capital lead to the maximisation of output and growth by widening markets and checking monopolies.

Criticisms of Supply-side Economics

The above prescriptions of supply-side economics have been criticised by economists on the following grounds:

1. Laffer Curve Controversial. The Laffer curve is an interesting but a controversial concept. No one knows with certainty either the location of the optimum point or the exact shape of this curve. The curve may peak at 40% or 90% tax rate, or it may peak in-between these rates. For instance, if we take the curve which peaks at point A in Figure 4, the present tax rate-T should be cut to T_1 to maximise revenue. On the other hand, if another curve peaks at point B, the tax rate T should be increased to T_2. Without the knowledge of either the peak or the shape of the curve, it is not possible to know the effect of reducing (or

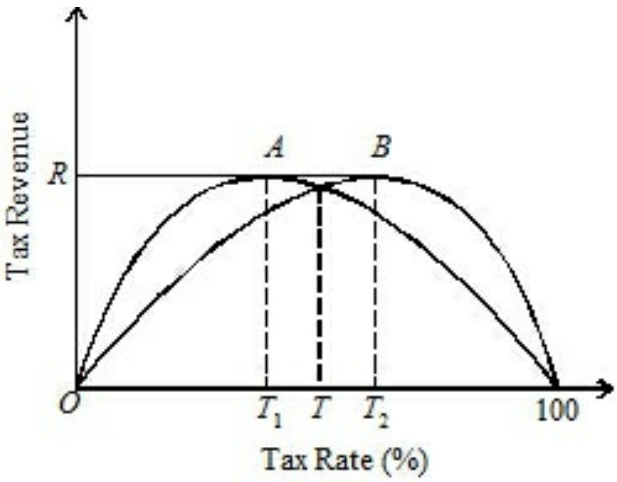


Fig. 4

increasing) the tax rate or tax revenue and economic activity. In fact, no body knows the exact shape of the Laffer curve or the relationship between tax rate and tax revenue.

---

2. For criticism of the Laffer curve, see the next section, and Fig. 4.

3. When the government spends more than it collects in taxes it finances its expenditure by selling its bonds. This is monetizing the deficit.

2. Tax Cuts do not bring High Growth Rate. Economists do not agree that cutting tax rates will lead to high growth rate and more tax revenue. They point out that high growth rate generates higher incomes which, in turn, generate higher tax revenue. Therefore, it is not reduction in tax rates that leads to the high growth rate of the economy.

3. Tax Cuts do not measure Work Effort. It is not possible to measure work effort specifically as a result of tax cut. No doubt, increased work effort leads to higher incomes and to increase in tax revenue.But the increased tax revenue may not be sufficient to compensate the government for the decrease in revenue due to the lower tax rate. Moreover, it is possible that people may work less when their disposable income increases with the lower tax rate.

4. Tax Cuts do not affect Target Incomes. Critics argue that some persons have 'target' real income. When taxes are reduced, they will work less and have more leisure to maintain their target income.

5. State Intervention Necessary. Supply-siders have been criticised for their policy of non-intervention by the state. But there are many contradictions in the working of the capitalist system which cannot sustain balanced growth of the economy. When the economy reaches full employment, a number of distortions and imbalances develop which fail to maintain full employment. Therefore, state intervention is necessary to remove them.

6. Supply-side Policies fail to bring Social Justice. Supply-side economists emphasise reduction in social spending, subsidies, grants and budget deficit with reduction in taxes. But such a policy has actually led to

huge budget deficits in the United States. Further, the policy of reducing social spending, subsidies and grants adversely affects the poor and unemployed and fails to bring social justice.

SELECTED READINGS

A. Laffer, "Supply-side Economics : A Symposium", *Economic Impact* Vol. 1 No., 37, 1982.

G. Gilder, "The New Arguments in Economics", *The American Review,* Vol. 26, No. 1,1981.

J. Tobin, "Supply Side Economics. What is it? Will it work?". *Economic Qutlook,* 1981.

EXERCISES

1. Explain the main features of Supply-side Economics.
2. Explain critically the Supply-side Economics.
3. Explain Supply-side Economics. What are its policy implications?
4. Explain critically the Laffer curve.

CHAPTER 56

THE NEW CLASSICAL MACROECONOMIC

INTRODUCTION

The new classical macroeconomics is an attempt to repudiate and modify Keynesian and monetarist views about the role of macroeconomic stabilisation policy in the light of the classical school of thought. The Keynesians advocate *demand management policies* both fiscal and monetary to stabilise the economy. They favour active interventionist fiscal and monetary policies. They do not regard the two policies as competitive but complementary to each other. But they depend more on expansionary fiscal policy to control recessions which threaten rising unemployment with little or no growth in the economy. However, they combine deflationary fiscal policy with monetary policy to control boom and inflation.

In contrast, monetarist hold that the economy is basically stable and when disturbed by some change in basic conditions will quickly revert to its long-run growth path. They are highly critical of discretionary fiscal and monetary policies. For such policies involve long and variable time lags which can make them ineffective and destabilising. However, they advocate an annual fixed percentage growth in money supply instead of discretion in monetary policy. Friedman believes that fiscal policy does not have any potent influence on the economy except that it affects the behaviour of money. Therefore, by setting and sticking to rules and not interfering, the government can follow a sound monetary policy in which there is maximum freedom for individual initiative and enterprise. The rules help to reduce people's expectations of inflation and thus create a

stable environment for investment and growth.

THE NEW CLASSICAL MACROECONOMICS

During the late 1970s when the debate between Keynesians and monetarists stalemated, the new classical macroeconomics emerged based on classical microeconomics. It was developed by Robert Lucas, Thomas Sargent, Robert Barro and Neil Wallace in America and Patrick Minford in England.

The new classical macroeconomics is based on the following principles or hypotheses:

(1) Markets Continuously Clear

(2) Rational Expectations

(3) Aggregate Supply Hypothesis

The hypotheses (1) and (3) are classical but their analysis is new. The second hypothesis on rational expectations is totally new. Therefore, these principles constitute the New Classical Macroeconomics which are discussed below.

1. Markets Continuously Clear

The new classical economists assume that all markets continuously clear in the economy. Prices and wages adjust instantaneously to clear markets. The economy is in a state of continuous equilibrium both in the short-run and long-run where all markets clear.

The new classicals differ from Keynesians and monetarists over market clearing. According to Keynesians, markets may not clear due to slow price adjustments. So the economy may remain in a state of disequilibrium. Monetarists assume that markets have a tendency to clear. Prices and wages are fairly flexible. Therefore, the economy may be in disequilibrium temporarily in the short run and attain equilibrium in the long run.

The new classicals assume that markets clear instantaneously and there is no disequilibrium even in the short run. Since price and wage adjustments are almost instantaneous, all unemployment is equilibrium unemployment. Whatever level of unemployment is found in the economy, it is the natural rate of unemployment or *voluntary unemployment.* An increase in the natural level of unemployment over time is the result of reluctance of people to take jobs due to lack of incentives.

Fig. 1 explains the new classical labour market equilibrium. Where SS_L. is the labour supply curve which is vertical (or inelastic) at ON_T labour force when wage rates are above the competitive level. DD_L is the labour demand curve. ON_T is the total labour force in the economy. The two curves intersect at E which is the market-clearing equilibrium point where ON workers are willing to work at the market wage rate OW. This is the full employment equilibrium. But NN_T (= EV) workers out of the total labour force (ON_T) are not prepared to work at the market wage rate OW. They are *voluntarily unemployed.* They may prefer a higher wage than the equilibrium rate or leisure or other activities, etc. to work.

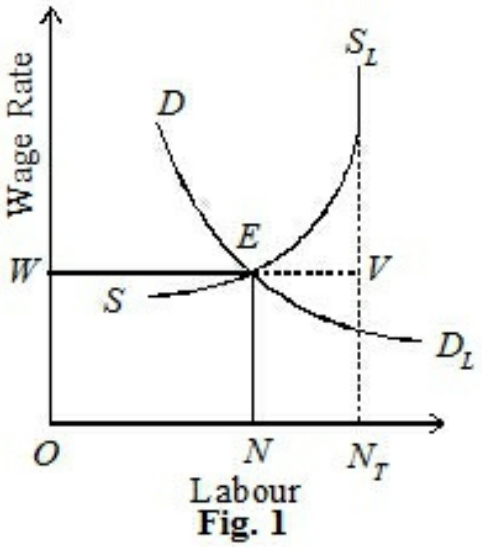


Fig. 1

2. Rationa l Expectations

One of the most important principles of the new classical macroeconomics is the rational expectations hypothesis. The Ratex hypothesis, as it is called, holds that economic agents (individuals, firms, etc.) form expectations of the future values of economic variables like prices, incomes, etc. by using all the economic information available to them.

The new classical economists use Ratex to explain the Phillips curve in the inflation theory. According to them, rational expectations are not based on past rates of inflation but on the current state of the economy and policies being followed by the government. Workers and firms base their

information on various forecasts made by specialists and agencies, and government announcements and reports. On the basis of such current information, they predict the rate of information.

Generally, such forecasts are wrong and what the government says is also not correct. So workers and firms base their expectations on imperfect information. It is thus on the basis of imperfect information that workers and firms make predictions which will frequently be incorrect. But such errors in predictions are random which make predictions about inflation either too low or too high. Any discrepancy between the actual and expected rate of inflation is only in the nature of random error.

Thus there is no possibility for the actual rate of unemployment to differ from the natural rate even temporarily. When people act rationally, they know that past increases in prices and the rate of change in prices have invariably been accompanied by equal proportional changes in the quantity of money. When people act on this knowledge, it leads to the conclusion that there is no trade-off between inflation and unemployment either in the short run or in the long run and the new classical Phillips curve is vertical at the equilibrium or natural rate of unemployment.

The new classical short-run vertical Phillips curve is shown in Fig. 2 as *PC* at the natural unemployment rate U_N. If people underpredict the rate of inflation (expected inflation rate is less than the actual rate), they will believe that aggregate demand has increased. As a result, output and employment rise. This shifts the short- run Phillips curve *PC* to the left as PC_1 because unemployment temporarily falls to U_1 below the natural rate U_N If, on the other hand, people overpredict the rate of inflation (expected inflation rate is more than the actual rate), they will believe that aggregate demand has fallen, and output and employment fall. This shifts the short-run Phillips carve *PC* to the right as PC_2 because unemployment

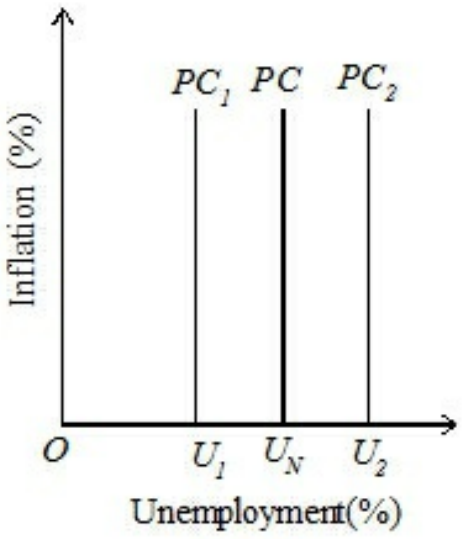


Fig. 2

temporarily rises to U_2, above the natural rate U_N. But the actual position of the short-run Phillips curve on the average will be PC at the natural unemployment rate U_N.

The new classical economists also explain the *downward sloping* short-run Phillips curve. Such a curve arises when people are not able to correctly predict about real wages. The new classical Phillips curve is vertical at the natural rate of unemployment shown as *PC* in Fig. 3. This is the *true* Phillips curve. To explain the downward sloping Phillips curve, called the *apparent* Phillips curve, we start at point *A* on the *PC* curve when the unemployment rate is 3% and the inflation rate is 4%. In order to reduce unemployment, the monetary authority unexpectedly increases the money supply to stimulate the economy. According to the Ratex hypothesis, firms have better information about prices in their own industry than about the general level of prices. They mistakenly think that the increase in prices is due to the increase in demand for their products. As a result, they employ more workers in order to increase output. Unemployment falls to 2%. The workers also mistake the rise in prices as related to their own industry. But wages rise as the demand for labour increases and workers think that the increase in money wages is an increase in real wages when the inflation rate rises to 6%. Thus the economy moves upward from point *A* to *B*. But soon workers and firms find that the increase in prices and wages is prevalent in most industries. Firms find that their costs have increased. Workers realise that their real wages have fallen due to rise in inflation rate to 6% and they press for increase in wages. But firms do not employ more workers. So the economy moves from point *B* to *A* which is the actual position of the short-run Phillips curve.

In such a. situation, workers *overpredict* the 4% rate of inflation. Employment will fall as workers believe that their real wages are lower than they actually are. So they work less. Output falls as firms believe that the relative prices of their products have fallen. With fall in employment and output, the economy moves from point *A* to *C* due to an *unanticipated* fall in wages and prices. Thus, points *B,A,C* trace out a downward-sloping *apparent* short-run Phillips curve PC_1 (in Fig. 3) of the new classical macroeconomics when people underpredict real wages and relative prices.

But the *true* short-run Phillips curve of the new classicals is always vertical like the *PC* curve.

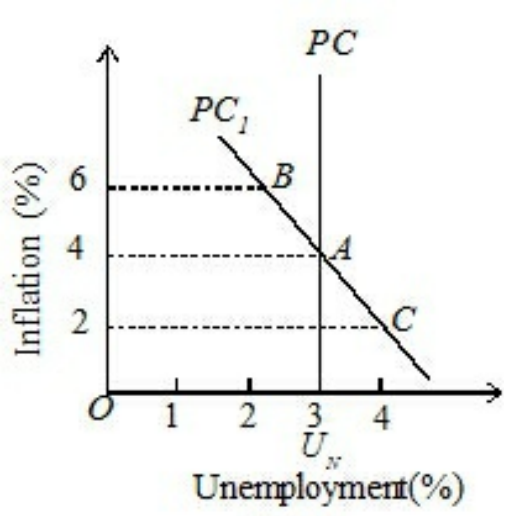


Fig. 3

3. Aggregate Supply Hypothesis

The new classical macroeconomics incorporates the Lucas[1] aggregate supply hypothesis based on two assumptions: (1) rational decisions taken by workers and firms reflect their optimising behaviour, and (2) the supply of labour by workers and output by firms depend upon relative prices. Thus the aggregate supply hypothesis is derived from optimising behaviour of workers and firms about supply of labour and goods which depend on relative prices only. We first study the labour market and then the goods market to explain the aggregate supply hypothesis.

---

1. Robert Lucas of the University of Chicago won Nobel Prize for Economics in 1995.

The Labour Market

Workers make decisions about work and leisure in the present with the future in mind. They also have some idea about the normal or expected real wage. If the current real wage is *above* the normal real wage, workers will have an incentive to work more in the present (take less leisure) in order to have more leisure (work less) in the future when the real wage is expected to be lower. On the other hand, if the current real wage is *less* than the normal real wage, workers will have an incentive to take more leisure (work less) in the present, in anticipation of working more in the future when the real wage is expected to be higher. This behaviour of workers to substitute current leisure for future leisure and vice versa is known as *intertemporal substitution.* From this, the new classical economists infer that the *short-run* supply curve of labour is r*elatively elastic* because expected changes in the real wage are temporary. But the *long-run* supply curve of labour is *vertical* because the real wage is permanent and the actual and expected price levels are the same.

In the new classical analysis, workers have *incomplete* information about price changes so that they mistake changes in general price level for relative changes in prices and thus change the supply of labour. This results from unanticipated shocks such as monetary disturbances which change aggregate demand.

The aggregate demand and supply analysis is used to illustrate the effects of unanticipated changes in aggregate demand on the real wage level and employment. In Fig. 4, $LRAS_L$ is the long-run aggregate supply curve of labour and $SRAS_L$ is the short-run supply curve of labour. *AD* is the aggregate demand curve. The labour market is initially in equilibrium at point *A* where the curves $LRAS_L$, $SRAS_L$ and *AD* intersect. Here the real wage rate *W*/*P* is fully anticipated and *OL* number of workers are employed. Suppose the monetary authority announces its intention to increase the money supply. This will have the effect of increasing aggregate demand. This shifts rightward the *AD* curve to AD_1. If the shift in aggregate demand is anticipated, rational agents will negotiate for higher real wage immediately on the basis of the expectation of rise in the price level. The $SRAS_L$ curve will shift upward to $SRAS_{L1}$. The real wage rate will move straight from *W/P* to W/P_2 on the vertical $LRAS_L$ curve and the labour market will move from *A* to *C* where the curves AD_1, $SRAS_{L1}$ and $LRAS_L$ intersect with no effect on the number of *OL* workers employed.

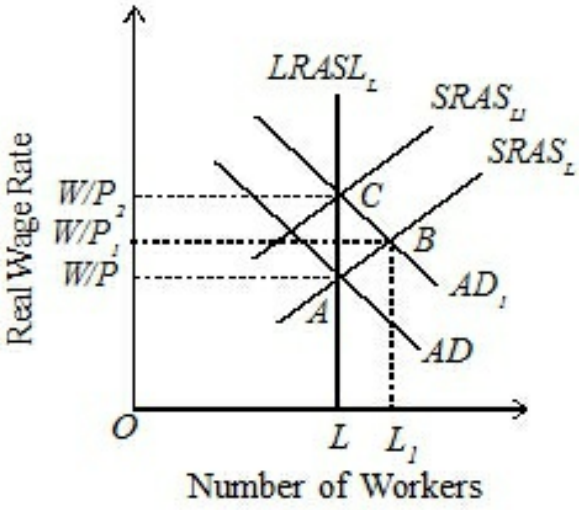


Fig. 4

If the shift in aggregate demand due to increase in the money supply is *unanticipated,* firms will misperceive the increase in general and relative prices. They will want to produce more and increase the demand for workers which will raise the real wage rate. In the figure, the *AD* curve will shift upward to AD_1 and intersect the $SRAS_L$ curve at point *B*. The number of employed workers will increase from *OL* to OL_1 alongwith the

rise in real wage to W/P_1. This increase in employment in the short-run is only temporary. But when firms fully adjust their price expectations in the long run, the $SRAS_L$ curve will shift to $SRAS_{LI}$ to intersect the AD_1 curve at *C* with no change in the level of *OL* workers employed, though at a higher real wage W/P_2.

The Goods Market

Consider the goods market in Fig. 5 where the economy is initially in equilibrium at point *A* where the curves *LRAS, AD* and *SRAS* intersect. Here the price level *OP* is fully anticipated and *OY* is the long-run equilibrium level of output. Suppose there is increase in aggregate demand due to *anticipated* increase in the money supply. This will shift the *AD* curve upward to the right to AD_1.

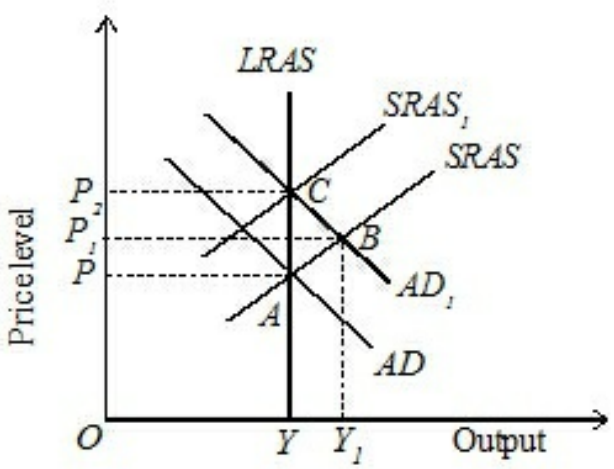


Fig. 5

As a result, there is an immediate upward revision of price expectations to OP_2. Firms increase the supplies of goods and the *SRAS* curve shifts upward to the left to $SRAS_1$. There is now a new equilibrium at point *C* where the curves AD_1, $SRAS_1$ and *LRAS* intersect. The price level moves straight from *OP* to OP_2 and the economy moves from *A* to *C* with no increase in the output level *OY*. However, if the increase in aggregate demand is *unanticipated* due to increase in the money supply, the economy moves from the initial equilibrium point *A* to *B* at the intersection of AD_1 and *SRAS* curves with the price level rising from *OP* to OP_1 and output increasing from *OY* to OY_1 level. But this will be only in the short run. When the economy goes through an adjustment process, it will return to its long-run equilibrium level of *OY* output at OP_2 price level.

Policy Implications of New Classical Macroeconomics

The new classical macroeconomics has a number of policy implications which are explained as under:

1. Policy Ineffectiveness Proposition. The new classical macroeconomic analysis holds that with rational expectations and flexible prices and wages, monetary policy, if anticipated in advance, will have no effect on output and employment in the short run. This is the *policy ineffectiveness proposition.* It is only an unanticipated increase in the money supply that will affect output and employment.

The policy ineffectiveness proposition is explained in Fig. 6 in terms of a supply curve of firms. The relative price at which firms sell the good is taken on the vertical axis and the quantity supplied on the horizontal axis.*SS* is the supply curve. OP_A is the anticipated relative price and OP_u is the unanticipated relative price of the good. Suppose the monetary authority increases the money supply and if prices are flexible, all prices will rise in the economy. If the increase in the money supply is *unanticipated*, firms think that their own prices have risen. They are fooled into thinking that the relative price of the good has increased from OP_A to OP_U. So they increase the quantity supplied from OQ to OQ_1. On the other hand, if the increase in the money supply is *anticipated*, firms cannot be fooled into thinking that relative price has increased. They know that prices of all firms have risen. So they keep their quantity supplied at OQ and there will be no change in output. Thus an anticipated increase in the money supply has no effect on output which proves the policy ineffectiveness proposition.

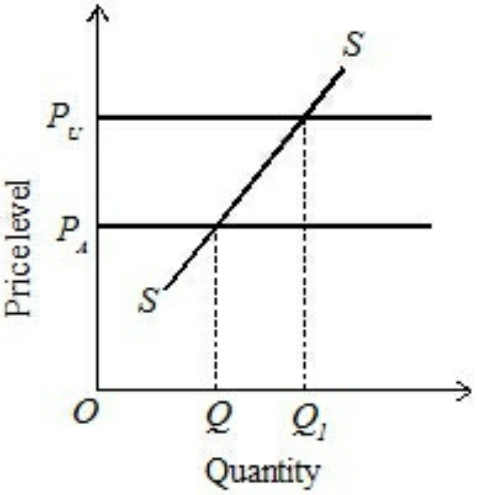


Fig. 6

2. Impotency of Systematic Monetary Policy. According to the new classical analysis, anticipated changes in aggregate demand will have no effect on output and employment even in the short run by pursuing a *systematic* monetary policy. A systematic monetary policy is one which takes into account any known “rule”. Such a policy can be fully predicted

by the private sector before the monetary authority actually acts upon it. So private buyers and sellers who anticipate increase in the money supply adjust their purchases and sales through flexible wages and prices. Further, the new classicists argue that *non-systematic* (or discretionary or unanticipated) monetary policy will only bring changes in output and employment around their natural levels. Therefore, to prevent unanticipated changes in aggregate demand and unemployment deviating from its natural level, the new classicals advocate clear monetary rules and avoidance of any discretionary monetary policy.

3. Policy Credibility. The new classical approach is based on the presumption that rational economic agents have expectations about what the monetary authority is going to announce and this influences their behaviour. But it is on the credibility of policy announcements of monetary authority that agents form expectations. Thus the new classical policy implies that announced (or anticipated) changes in monetary policy will have no effect on output and employment even in the short run provided the policy is *credible.* Suppose there is an announced and credible reduction in the money supply. This will immediately lead to a downward revision of inflation expectations of rational economic agents. This will, in turn, enable the monetary authority to have disinflation without output and employment costs.

4. The Lucas Critique. Robert Lucas criticised the building of econometric macroeconomic models of the economy for policy evaluation. According to Lucas, such models were based on parameters derived from past data collected under particular policies. Any attempt to use such macroeconomic models to predict the consequences of alternative policies may be wrong. This is because the parameters of such models may change as economic agents adjust their expectations and behaviour to the new policy.

Lucas argued that although economic agents act in a certain way, it is erroneous to assume that they would continue to act in the same way, if economic policy is changed. Suppose workers anticipate inflation to be 5 per cent next year and they demand 5 per cent wage increase. Anticipating it, if the monetary authority increases the money supply, inflation rises to

10 per cent. This reduces the real income of workers, and firms finding cheap labour, employ more workers to make more goods. This would increase output by reducing the real wage of workers whose expectations of 5 per cent turn out to be wrong. According to Lucas, such a policy may succeed once or twice. But if the monetary authority continues such a policy, people would expect higher inflation in future and the policy would fail. The monetary authority cannot fool the people all the time.

Thus the Lucas critique points out that workers and firms are assumed to choose their actions in the light of existing policies. If there is a major change in policy, it will change people's behaviour and expectations. The general implication of the Lucas critique is that the effects of policy changes are difficult to forecast accurately and they can be learnt by experience.

5. Policies to Increase Aggregate Supply. One of the important policy inplications of new classical macroeconomics relates to the nature of policies to be followed by the authorities to increase output and reduce unemployment. In the new classical analysis, changes in output and employment are based on the equilibrium supply decisions of firms and workers given their perceptions of relative prices. It follows that the appropriate policy measures to increase output and reduce unemployment are directed toward increasing aggregate supply of output and labour. New classical macroeconomists recommend a variety of measures to increase output and reduce unemployment that indirectly increase aggregate supply of output and labour. They relate to reduction in the power of trade unions, reduction in unemployment benefits, tax reforms to remove poverty and raise incomes of the unprivileged, measures to increase geographical and occupational mobility of labour, etc.

Criticisms of New Classical Macroeconomics

The new classical macroeconomics has been criticised mainly on the basis of its hypotheses and policy implications.

1. Rational Expectations Hypothesis Unrealistic. The rational expectations hypothesis which is the backbone of the new classical approach has *four* main objections. *First,* it costs much to acquire, process

and disseminate publicly available information. So the majority of economic agents cannot act on the basis of rational expectations. *Second,* the critics point out that information available to government differs from that available to firms and workers. Consequently, expectations of the latter about the expected rate of inflation need not necessarily diverge from the actual rate only by the random error. But the government can accurately forecast about the difference between the expected inflation rate and actual rate on the basis of information available with it.*Third,* event if both people and government have equal access to available information there is no guarantee that expectations will be rational. *Fourth,* as the cost of acquiring, processing and disseminating publicly available information is very high, economic agents may form expectations which are systematically wrong. Thus the rational expectations hypothesis is unrealistic and the new classical macroeconomics which is based on it stands on weak foundations.

2. Markets do not Continuously Clear. Critics do not accept the hypothesis that all markets continuously clear. They point out that prices and. wages are not flexible. There is collective bargaining in the labour market which leads to wage contracts leading to stickiness of money wages. The rigidity of wage rates implies that they adjust to market forces relatively slowly because wage contracts are binding for two or three years at a time. Similarly the expected price level at the beginning of the period is expected to hold till the end of the period. As a result, labour market and goods market are unable to clear continuously. As pointed out by Tobin, “The market-clearing assumption is just an assumption and nothing more than that”.

3. Aggregate Supply Hypothesis Unacceptable. Economists do not accept the aggregate supply hypothesis that changes in output and employment reflect voluntary response of workers and firms to perceived changes in relative prices. According to them, it is changes in aggregate demand announced by the monetary authority that influence output and employment both in the short run and long run.

4. Policy Implications Unacceptable. Critics do not accept the policy implications of new classical macroeconomics because they are derived

from unrealistic hypotheses. Economists like Phillips, Taylor and Fischer have demonstrated that if wages and prices are not completely flexible, monetary policy becomes *effective* in the short run. It can influence output and employment in the short run even if expectations are rational.

Further, as firms do not know enough about the structure of the market to estimate the market-clearing price level and there are non-clearing labour markets due to wage rigidity, economists do not accept the impotency of monetary policy.

Empirical Evidence

There has been some empirical evidence for and against the new classical macroeconomics. Economists like Sargent, Minford, Barro, Gordon, Blinder, etc. have constructed econometric models to test the hypotheses and policy implications of new classical macroeconomics. The results of main empirical evidences are as under:

1. Empirical evidence on European depression shows microeconomic interferences in labour markets in the form of generous unemployment insurance when unemployment was extremely low in 1973.

2. Empirical research has not been able to find large intertemporal substitution effects in labour market

3. Lucas in his model of 1973 found evidence in support of the new classical Phillips curve that it was vertical in the short run. But Gordon's econometric study for Europe in 1987 concluded that the original empirical Phillips curve existed.

4. A number of empirical studies, one by Muth himself in 1985, have questioned the validity of rational expectations hypothesis. They used directly observed data on expectations to test rationality. These tests rejected the rational expectations.

5. Rotemberg statistically tested some macroeconomic models of rational expectations in 1984 on the basis of the three hypotheses viz, expectations are rational, markets continuously clear and aggregate supply, of the new

classical theory. When tested jointly, the joint hypothesis was rejected.

6. Barro in his statistical test of unanticipated changes in money growth on output and employment came to the conclusion that it is unanticipated changes in the money stock rather actual money growth that affect output and employment with quite long lags of two to four years.

SELECTED READINGS

R.J. Gordon, *Macroeconomics,* 6/e, 1993.

A. Klamer, *The New Classical Macroeconomics, 1984.*

N.G. Mankiw, *Macroeconomics,* 2/e, 1994.

H.R. Vane and J.L. Thomson, *An Introduction to Macroeconomic Policy,* 1985.

EXERCISES

1. Explain the main tenets of the New Classical Macroeconomics. Discuss its policy implications.

   Explain the hypotheses on which the New Classical Macroeconomics is based. On what grounds has this approach been criticised ?

CHAPTER 57

THE REAL BUSINESS CYCLE THEORY

INTRODUCTION

The real business cycle theory has been evolved out of the American new classical school of 1980s. It is the outcome of research mainly by Kydland and Prescott[1], Barro and King,[2] Long and Plosser[3] , and Prescott[4] . Later, Plosser[5] , Summers, Mankiw and many other economists gave their views of the real business cycles. They view aggregate economic variables as the outcomes of the decisions made by many economic agents acting to maximize their utility subject to production possibilities and resource constraints. Their views mainly relate to technology shocks, labour market, interest rate, role of money, fiscal policy, prices and wages in business cycles. They are explained below.

---

1 F.E. Kydland and E.C. Prescott, "Time to Build and Aggregate Fluctuations," Econometrica, Nov. 1982.

2 R.J. Barro and R.G. King, "Time-separable Preferences and Intertemporal-Substitution Models of Business Cyles," Working Paper 888, N.B.E.R., 1982.

3 J.B.Long and C.I. Plosser, "Real Business Cycles, "J.P.E. Feb., 1983

4 E.E. Prescott, "Theory Ahead of Business Cycle Measurement, "Federal Reserve Bank of Minneapolis Quarterly Review, Fall, 1986.

5 Charles I. Plosser, "Understanding Real Business Cycles, "Journal of Economic Perspectives, 3, Summer 1989.

N.B. Kydland and Prescott were awarded Nobel prize in Economics in 2004.

ROLE OF TECHNOLOGICAL SHOCKS

The theory of real business cycles explains short-run economic fluctuations based on the assumptions of the classical theory. According to this theory, business cycles are the natural and efficient response of the economy to economic environment. They are primarily caused by real or supply side shocks that involve exogenous large random changes in technology. An initial shock in the form of a technological advance shifts the production function upward. This leads to increase in available resources, investment, consumption and real output. With the increase in investment, the capital stock increases which further increases real output, consumption and investment. This process of expansion of the economy continues erratically due to changes in technology over time. According to Plosser, "It is a purely real model, driven by technology disturbances, and hence, it has been labeled a *real business cycle model*."

Assumptions

The real business cycle thoery is based on the following assumptions :

1. There is a single commodity in the economy.

2. Prices and wages are flexible.

3. Money supply and price level do not influence real variabels such as output and employment.

4. Fluctuations in employment are voluntary.

5. Population is given. So there is fixed labour force.

6. There are rational identical economic agents in the economy.

7. These agents make optimising decisions.

8. Every one has the same preferences which depend only on consumption in each year.

7. More consumption is preferred to less so that the marginal utility from

consumption diminishes.

8. The economy is subject to irregular (random) real supply side shocks.

9. It is a single sector economy.

10. There are substantial changes in the rate of technology that affect the whole economy (which is viewed as a single sector).

11. There is constant return to scale production-technology.

12. The economy is in a steady state.

Technological Shock

Given these assumptions, the production function of the economy is given by

$$Y = Zf(K,N)$$

where Y is total output, Z is the state of technology, K is predetermined capital stock and N is labour input. The produced output can either be consumed or invested.

Assuming that population is given and there is a fixed labour force, output depends on technology and capital stock. So output is determined by the production function, $Y = Zf(K)$. The capital stock, K depreciates at the rate S, so that the undepreciated capital stock evolves as $(1-\delta)K$. This capital stock is available as input for production in the next period.

With a capital stock K, output is Y and the total resources available in the economy in the current period are $Y + (1-\delta)K$.

Since $Y = Zf(K)$, the total resources can be expressed as $Zf(K) + (1-\delta)K$. These resources can either be consumed or accumulated as capital to be used as investment for the next period.

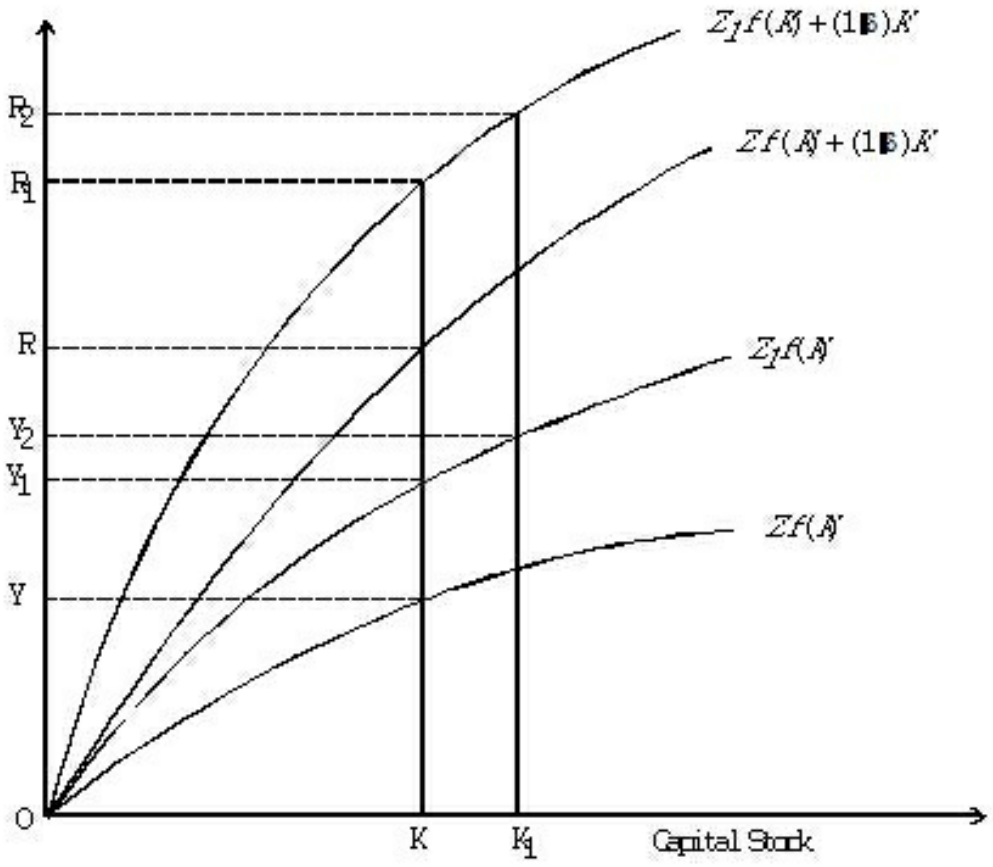


Fig. 1

A real business cycle is generated in a steady state economy when there is a positive exogenous and permanent technological shock. This leads to increase in productivity. As a result, the aggregate production function shifts upward. The improvement in technology from the initial level Z to Z_1 and the consequent upward shift of the production function from $Zf(K)$ to $Z_1 f(K)$ is shown in Figure 1. Given the initial capital stock OK, output increases from OY to OY_1. As a result, total resources increase from OR to OR_1 and the total resources curve shifts upward from $Zf(K)+(1-\delta)K$ to $Z_1 f(K)+(1-\delta)K$. With the increase in total resources, both current consumption and capital accumulation also increase. There is increase in capital stock to OK_1. With no change in technology, the increase in capital stock to K_1 in the next period leads to a further rise in output to OY_2 and the increase in total resources to OR_2. In this way, the economy continues to expand when consumption, investment and output increase gradually leading to a new steady state.

But the path to a new steady state will not be smooth. With a permanent technological advance, consumption and investment increase in the next period. But the increase in total resources and output is smaller than in the initial period. In Figure 1, $R_1R_2 > R\,R_1$ and $Y_1Y_2 < YY_1$. In the long run, there is a gradual decline in investment and consumption even when output continues to increase at a decreasing rate till the economy reaches the new steady state. The paths of this real business cycle are illustrated in Figure 2.

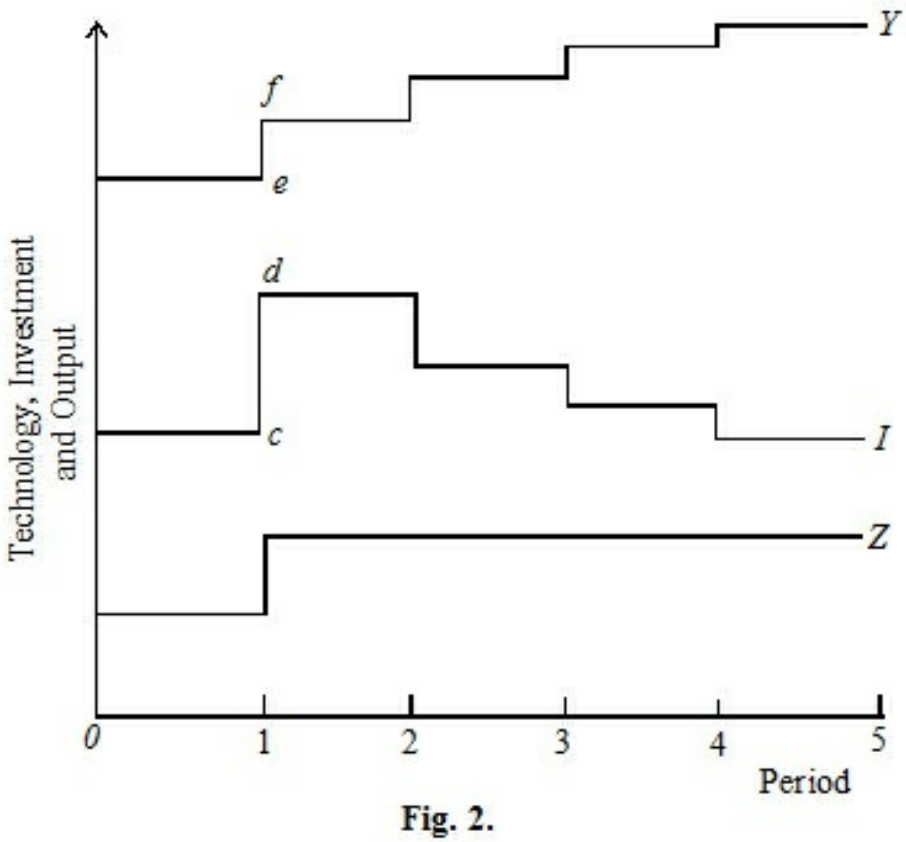


Fig. 2.

In period 1, there is a permanent technology shock which advances technology *Z* from *a* to *b* . This leads to increase in investment *I* from *c* to *d* and output *Y* from *e* to *f* . Given the same level of technology *Z*, shown as the horizontal curve, the investment curve *I* gradually falls in subsequent periods but the output curve *Y* continues to increase at a decreasing rate till the economy reaches the new steady state in period 5.

A *recession* in the real business theory is just the reverse of the expansion. A shock of decline in technology reduces *Z* and shifts the production function downward and decreases the available resources. This starts a process of decline in investment, consumption, output and employment. But the models of real business cycle do not explain a recession.

Labour Market

The real business cycle theory emphasises that there is intertemporal substitution of labour in the labour market. When a technology advance leads to a boom, the marginal product of labour increases.There is increase in employment and real wage. In response to a high real wage, workers reduce leisure. On the contrary, when technology is unfavourable and declines, the marginal product of labour, employment and real wage rate are low. In response to a low real wage, workers increase leisure. Thus an important implication of real business theory is that the real wage is procyclical.

Interest Rate

The real business cycle theory also takes into account the role of real interest rate in response to a technological shock. The real interest is equal to the marginal product of capital. When a favourable technological change leads to a boom, the marginal product of capital and the real interest rate rise. On the contrary, an unfavourable technical change leading to a recession reduces the marginal product of capital and the real interest rate. When the economy reaches the new steady state, the real interest rate eventually returns to its initial level.

Flexibility of Wages and Prices

The real business cycle theory assumes than wages and prices are flexible. They adjust quickly to clear the markets. There are no market imperfections. It is the "invisible hand" that clears the market and leads to an optimal allocation of resources in the economy.

Neutrality of Money

Money plays no role in the real business cycle theory. Money is neutral. It is a veil. Money does not affect such real variables as employment and output. The role of money is to determine the price level. The money supply is endogenous in the real business cycle theory. It is fluctuations in output that cause fluctuations in the money supply. For instance, when there is a favourable technological change, the output increases and the

quantity of money demanded rises. The banking system responds by advancing more loans and the central bank increases the money supply. With the money supply increasing, prices rise.

Fiscal Policy

Fiscal policy has little role to play in the real business cycle theory. Since the "invisible hand" guides the economy, the government role is limited. In fact, business cycles are the natural and efficient response of the economy to favourable and unfavourable technological shocks. A fiscal policy measure such as a tax on income will adversely affect output and employment. An individual may choose more leisure to work leading to reduction in consumption, investment and output. To avoid tax distortions and meet its requirements, the central bank increases the money supply in the economy. So the government has no role in stabilisation policy.

Criticisms of the Real Business Cycle Theory

The real business cycle theory has been highly controversial. Its exponents maintain that it is a realistic theory based on large fluctuations in output and employment displayed by the US economy. But its foremost critic, Lawrence Summers points out that the real business cycle models "have nothing to do with the business cycle phenomena observed in the United States or other capitalist economies." According to Mankiw, "Real business cycle theory does not provide an empirically plausible explanation of economic fluctuations." Summers[6], Mankiw[7] and many other economists have criticised the real business cycle theory on the following grounds.

1. Technological Shocks. In the real business cycle theory as put forward by Prescott, the only driving force behind cyclical fluctuations is technological shocks. Critics do not agree with this. According to them, technological shocks leading to changes in total factor productivity are hard to find. There is also no direct evidence of the existence of large technological shocks. Therefore, the existence of large changes in technology is an unjustified assumption of real business cycle theory.

6 Lawrence H.Summers, "Some sketical observations on Real Business Cycle Theory," Federal Reserve Bank of Minneapolis Quarterly Review, 10 Fall, 1986.

7 N. Gregory Mankiw, "Real Business cycles : A New Keynesian Perspective," Journal of Economic Perspectives, 3, Summer, 1989.

2. Other Factors. This theory takes into account only supply side factors and ignores other factors like change in demand that cause business cycles.

3. Intertemporal Substitution. In the real business cycle theory, there is intertemporal substitution of labour and work. Over the business cycle, individuals reduce the supply of labour in response to small reductions in the real wage or small decreases in real interest rate. If individuals expect increases in their real wage, they will enjoy leisure today and work more in the future. If they expect decreases in their real wage, they will work hard today and enjoy leisure in the future. But it is unlikely that individuals are so responsive to intertemporal changes in real wage. Empirical studies of individual labour supply overtime have shown that expected changes in real wage lead to only small changes in hours worked. Thus individuals do not respond to expected real wage changes by reallocating leisure to work overtime.

4. Voluntary Employment. In the real business cycle theory, fluctuations in the employment are assumed voluntary. So it does not consider unemployment. When large number of people are searching for work in a recession, they are unable to find it. The theory explains why employment falls. But its explanation is based on reductions in the real wage and the real interest rate which send signals to workers that there is no work available in the labour market. Thus there is no scope for unemployed workers to find jobs and the labour market does not clear in a recession.

5. Exchange Mechanism. According to Summers, a fundamental objection of Prescott's real business theory is that it ignores breakdown of the exchange mechanism during depressions in cyclical fluctuations.

6. Neutrality of Money. The real business cycle theory assumes neutrality of money. But according to critics, the empirical evidence does not support that money is neutral in the short run. They point out that

money does affect such real variables as output and employment in a boom and a recession. When there is increase in money growth and inflation, output and employment are high in a boom, and vice versa in a recession.

7. Wages and Prices. This theory assumes that wages and prices are flexible. But critics point out that wages and prices are inflexible. They believe that changes in monetary policy lead to short-run aggregate demand disturbances that can have important real effects on output and employment because of the nominal price and wage rigidities.

8. Fiscal Policy. The real business cycle theory assumes that the government has no role to play in stabilisation policy. If the government adopts policies to stabilise employment, they are ineffective and can harm the economy by impeding the invisible hand. But crities do no agree that stabilisation policy has no role to play.

9. Negative Technological Shocks. This theory does not explain large negative technological shocks that mark recession. Historical evidence shows that periods of disinflation are of low output and unemployment. These effects are completely absent in the real business cycle theory.

10. Incomplete Theory. The real business cycle theory is an imperfect and incomplete theory. It does not explain the turning points of the business cycle. No doubt real supply shocks have important effects on output and employment, they do not create peaks and troughs in the business cycle as actually observed.

Conclusion. Despite these criticisms, as oberved by Mankiw, "The real business cycle theory has served the important function of stimulating and provoking the scientific debate, *but it* will ultimately be discarded as an explanation of observed fluctuations."

SELECTED READINGS

A.B.Abel and B.S. Bernake, Macroeconomics, 2/e 1995

R.T. Froyen, Macroeconomics : Theories and Policies, 7/e, 2002.

R.E. Hall and J.B. Taylor, Macroeconomics, 5/e, 1997.

R.E.A. Farmer, Macroeconomics, 2/e, 2002.

EXERCISES

1. Critically explain the real business cycle theory.
2. "It is a purely real model, driven by technology disturbances" (Plosser). Critically examine his view in the light of real business theory.

CHAPTER 58

NEW KEYNESIAN ECONOMICS

INTRODUCTION

New Keynesian economics is the school of thought in modern macroeconomics that differs from the traditional Keynesian principles and adjusts and reformulates them in response to their criticisms of the new classical economists. Some of the leading new Keynesian economists are George Mankiw, Robert Gordon, David Romer, Oliver Blanchard, Lawrence Summers, Joseph Stiglitz and Bruce Greenwald.

DIFFERENCES BETWEEN NEW CLASSICAL AND NEW KEYNESIAN MACROECONOMICS

New classical and new Keynesian economists differ over the following issues :

1. New classical economists argued that Keynesian economics was theoretically inadequate because it was not based on microeconomic foundations. According to them, macroeconomic models should be based on firm microeconomic foundations. New Keynesians agree on this but they differ how markets work. New classical economists base their models on perfectly competitive consumer, producer and labour markets. On the other hand, new Keynesians base their models on the real world imperfectly competitive markets where consumers, producers and labour market participants operate with imperfect information.

2. New classicals base their theories on market-clearing models where

demand and supply adjust quickly on the assumption that wages and prices are flexible. New Keynesians believe that market-clearing models cannot explain short-run economic fluctuations. So they base their models on sticky wages and prices that also explain why involuntary unemployment exists.

3. New Keynesian economics differs from new classical economics in explaining aggregate fluctuations in terms of microeconomic foundations. The new classicals explain the forces at work in terms of rational choices made by households and firms. But in new Keynesian analysis, households and firms do not coordinate their choices without costs. And cordination costs lead to coordination failure.

4. New classicals and new Keynesians also differ over the notion of equilibrium. In new classical models, markets clear continuously and wages and prices adjust quickly so that the quantity of labour demanded equals the quantity of labour supplied and there is full employment equilibrium. But in new Keynesian models, wages and prices fail to adjust rapidly enough to clear markets within a short time so as to keep the quantity demanded of labour equal to its quantity supplied. But this is an unemployment equilibrium. Economists call it disequilibrium or low-employment equilibrium. In fact, in new Keynesian economics, the actual quantities of labour demanded and supplied do not balance but the *expected* quantities of labour demanded and supplied balance.

5. New classicals and new Keynesians differ substantially over the use of stabilisation policy. The new classical analysis holds that with rational expectations and flexible prices and wages, and anticipated changes in aggregate demand will have no effect on output and employment in the short run by following a systematic monetary policy. This is the *policy ineffective proposition*. Therefore, new classical economists advocate monetary rules and avoidance of discretionary monetary policy to prevent unanticipated changes in aggregate demand when unemployment deviates from the natural level. In new Keynesian economics, when there is decrease in aggregate demand due to wage and price rigidities and market failures, active monetary and fiscal policies can prevent fall in output and employment.

MAIN FEATURES OF NEW KEYNESIAN ECONOMICS

New Keynesian economics was conceived in the late 1970s but several strands have evolved in new Keynesian macroeconomic theories/models since the mid 1980s. Some of the importands strands are discussed under four broad headings :

1. Sticky nominal (money) wages

2. Sticky nominal prices

3. Sticky real wages

4. Coordination failures

1. STICKY NOMINAL WAGES

In the classical theory of labour market, there is always full employment in the economy and no involuntary unemployment. In case of unemployment, a cut in money wages can achieve full employment. Firms can instantly adjust the quantities of labour they employ without any cost because of the flexibility of money wages. In the Keynesian theory, involuntary unemployment exists which can be removed by cut in real wages by increasing aggregate demand, output and employment. Keynes held that money wages are sticky. Within the Keynesian tradition, new Keynesian economists have developed the new Keynesian theory of the labour market based on nominal wages stickiness.

Assumptions

This theory is based on the following assumptions:

1. Nominal wages are sticky in the labour market.

2. They are set on the basis of contracts for a stipulated period.

3. They are set to make the expected quanity of labour demanded equal

the expected quantity of labour supplied.

4. Trade unions and firms form a rational expectation of future demand and supply of labour.

5. They agree on a wage that makes the expected quantity of labour demanded equal to the expected quantity supplied on the average over the period of contract.

6. Firms determine the employment level.

7. Workers are prepared to supply the required quantity of labour demanded at the fixed money wage over the period of the contract.

8. The employment level is determined by the actual demand for labour.

Explanation

Given these assumptions, in th new Keynesian theory of the labour market, money wages are set in contracts between workers (unions) and employers (firms) which maintain money wages over an agreed period. Such contracts are made because the demand for labour increases during recoveries and decreases during recessions which require changes in wage rates. Therefore, workers and employers find such wage contracts advantageous because there are high costs of collecting information about labour and negotiating frequent wage agreements.

When unions and firms start negotiations for agreed money wage rates over a stipulated period, they take a view of the expected demand for and supply of labour on the average. They know that setting very high wages rates will lead to very low employment on the average and large unemployment. On the other hand, setting very low wage rates will lead to a shortage of labour. Thus both parties act rationally and agree on such wage rates that make the expected quantity of labour demanded equal to the expected quantity of labour supplied.

The determination of money wage rate in the new Keynesian theory of labour market is shown in Figure 1. During the period of the contract, the

effective supply of labour is the horizontal line W_0W which shows the agreed wage rate of OW_0. Workers agree to supply the quantity of labour which the firms demand at this contracted wage rate. In the figure, S is the expected supply curve of labour and D_0 is the expected demand curve for labour. These curves meet at point E on the agreed money wage rate OW_0 where the expected labour employed is OQ_0. If the demand for labour happens to be more than expected at D_2, the quantity of labour employed increases to OQ_2. If the demand for labour happens to be D_1, the labour employed falls to OQ_1.

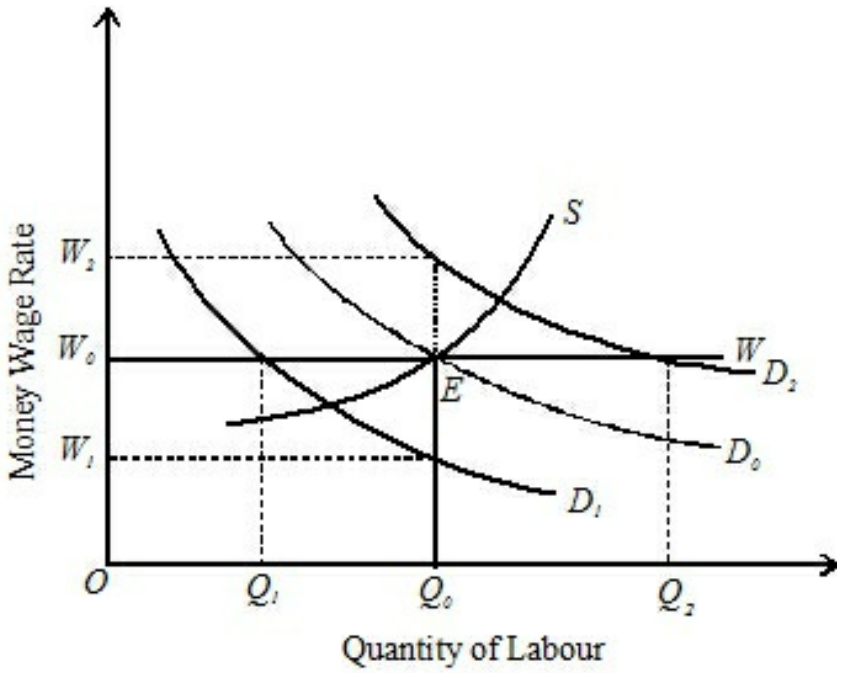


Fig. 1

The above analysis shows that the quantity of labour employed depends on the expected demand for labour. The expected demand for labour is determined by the expected price level and expected forecasts about the marginal product of labour (MP_L). MP_L, in turn, determines the quantity of labour which the firms will employ at each possible real wage rate. Similarly the expected supply of labour is also based on the expected price level and on expectations about the number of workers available for work at different real wage rates.

Suppose the price level rises or the marginal product of labour increases. These increases will shift the demand for labour curve to the right from D_0 to D_2 and equivalently the money wage rate that firms are willing to pay for that employment level. So the quantity of labour employed increases

to OQ_2 from OQ_0 and money wage rate from OW_0 to OW_2. In the opposite case with a fall in the price level or decrease in the marginal product of labour, the demand for labour, curve will shift the to left from D_0 to D_1 and equivalently the money wage rate will be reduced from OW_O to OW_1. The firms will reduce the quantity of labour employed from OQ_0 to OQ_1.

In the above analysis, when the price level rises or falls by say 50%, the money wage rate also declines or increases by 50%, the agreed money wage rate OW_0 remains the same. It is only when the demand for labour happens to be the same as expected at D_0 that the level of employment is equal to its expected level of OQ_0 . However, under gradually increasing wage rates, workers who stick to their jobs for a long time in the same firm get less than the value of their marginal product as they approach retirement. But over the long period, they are paid wages on the average equal to the value of their marginal product.

Staggered Wage Contracts Theory

One of the theories of nominal wage rigidities in the new Keynesian analysis is of staggered contracts. In the staggered contracts approach, not all labour unions sign contracts at the same time. There is no syncronization of contracts. Contracts are renewed at various times so that the dates on which new contracts start are staggered and they overlap. Such overlapping long-term wage contracts lead to nominal wage rigidity. During the contract period, the wage rate is fixed and linked to the expected rate of inflation by price indexing.

If prices are expected too high, workers will demand larger wage increases and firms will be willing to pay them because their own prices are expected to rise. Besides expectations of inflation, wage fixation is influenced by expectations of wages paid to other workers and the level of employment.

Taylor[1] in his staggered contract approach takes aggregate nominal demand indexation as a source of nominal wage rigidity. He assumes the nominal wage fixed over the period of the contract at a level that depends

on the expected price and expected future demand and output. A monetary disturbance affects demand and output during the period of contract until a new contract is negotiated.

Suppose the monetary authorities reduce the money supply in the economy which reduces aggregte demand and output. This requires a proportionate adjustment in nominal wages to maintain full employment. Since the wage contracts are staggered, wage adjustment is very slow in response to changes in aggregate demand and output. This makes nominal wages sticky.

2. Sticky Nominal Prices : Menu Costs Hypothesis

The classical and new classical microeconomic theories are based on the assumption of flexibility of prices where prices clear markets by adjusting demand and supply quickly. New Keynesian economists, on the other hand, believe in the stickiness of prices in the short-run. Markets do not clear quickly because adjusting prices is costly. Price adjustments involve costs to firms for frequently adjusting prices of their goods. A large sector of the economy is made up of price-makers who sell goods in monopolistic or imperfectly competitive markets. For them, adjusting prices is costly

The costs of adjusting prices are called the menu costs. Changing prices requires the use of resources by a firm. It has to print new price lists (menus), catalogues, and other printed material. A super market has to relabel all products and shelves with the new prices. A hotel and a restaurant has to reprint its menu with new prices. Meetings, phone calls, and trips by representatives of a firm to renegotiate with suppliers, all fall under the category of menu costs.

In the menu costs approach to sticky prices, it is profitable for firms to react to small changes in demand by keeping prices constant over a short period and responding with changes in output. Because of menu costs, firms do not change their prices every time with a change in demand conditions. Menu costs are incurred each time prices are changed periodically rather than continuously. Thus menu costs explain the short-

run stickiness of prices.

In the menu cost hypothesis, prices adjust slowly because changes in prices have externalities. When one firm reduces the price of a product, it benefits other firms in the economy. When it reduces the prices it charges, it lowers the average price level slightly and thereby raises real income. The increase in real income, in turn, raises the demand for the products of all firms. This macroeconomic impact of one firm's price adjustment on the demand for the products of all other firms is called an *aggregate-demand externality* by Mankiw. With aggregate demand externality, small menu costs can make prices sticky.

Assumptions

The sticky nominal price analysis of menu costs is based on the following assumptions :

1. There is an imperfectly competitive market which consists of a number of monopolistic competitive firms.d

2. Firms produce standardized or differentiated products.

3. Firms are price-makers having some control over the prices of their products.

4. Price adjustments involve costs for firms.

5. The demand curves are linear.

6. The marginal cost curve is horizontal.

---

1. J.B. Taylor, "Staggered Wage Setting in a Macro Model," *A.E.R.*, May 1979.

Explanation

Given these assumptions, we illustrate in Figure 2 how menu costs work and affect the adjustment of price and quantity of a firm's product. Let us suppose that the firm's demand has declined so that the original demand

curve D_0 shifts to the left to D_1 and correspondingly its original MR_0 curve to MR_1. Similarly its marginal cost has also declined. It has been shown as MC_1 which remains fixed. The original marginal cost curve MC_0 has not been shown to simplify the figure. The original price is OP_0 and quantity is OQ_0 when MR_0 intersects MC_1 at E. The firm's profit is $KEAP_0$. With the decline in demand, the intersection of MR_1 and MC_1 is at F and the price falls to OP_1 and quantity to OQ_1. Consequently, profit declines to $KFCP_1$. If the menu costs are high, the firm will keep the price at OP_0, reduce output to OQ_2 and earn $KGBP_0$ profit. The firm will reduce price to OP_2 only if the extra profit ($KEDP_2$–$KGBP_0$) exceeds the menu cost. The firm will not, therefore, reduce the price and there will be nominal price rigidity at OP_0.

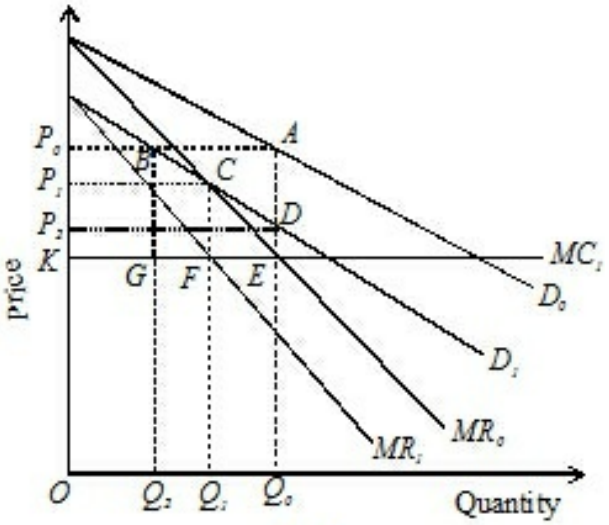


Fig. 2.

Criticisms

The menu costs approach has been criticised on the following grounds :

1. The menu costs approach is defective in that it considers only costs of price adjustment and not costs of output adjustment.

2. This approach assumes that marginal cost moves in proportion with demand. As demand rises or falls, marginal cost also increases or declines in the same proportion. In fact, no firm can assume that its marginal cost will be perfectly correlated with its aggregate demand.

3. This hypothesis tries to explain nominal rigidities in adjustments of the level of prices. But it fails to explain rigidities in adjustments of the *rate of change* of prices.

4. Critics point out that menu costs are small and have become smaller as computers allow the printing of menus at a small marginal cost.

5. Economists do not agree that menu costs can explain price stickiness in the short run because they are very small. Small menu costs cannot explain recession in the economy.

6. Another flaw is that small menu costs may be important for an individual firm but they are unlikely to affect the economy as a whole.

3. Sticky Real Wages

In the new classical labour theory, labour market is cleared continuously at the market-clearing real wage rate but it does not explain involuntary unemployment. On the other hand, the new Keynesian theories focus on the real wage rigidity where workers are not paid market-clearing wage and involuntary unemployment exists even in the long run. There are four main approaches to real wage rigidities. They are (a) Asymmetric information model, (b) implicit contract theory, (c) insider-outsider theory, and (d) efficiency wage theory.

(A) Asymmetric Information Model

Asymmetric information is a situation in which some persons have more information about some things than others. This idea that each person has asymmetrical information relative to others was used to develop a labour market model by Grossman and Hart. They asumed that managers know more about the interests of the firm than do the workers. Given this better knowledge, it is possible and profitable for managers to deceive the workers about the real position of the firm. They enter into contracts with workers and employment commitments whereby the firm pays them rigid real wages. However, there is an employment commitment in this model that tends to increase the amount of employment in the firm.

(B) Implicit Contract Theory

Two American economists, Baily and Azariades, have developed the

implicit contract theory. Usually employment contracts between workers and firms are explicit agreements. But often there are other dimensions that are not written in the actual contracts. These dimensions are called implicit contracts. Workers and firms enter into implicit contracts concerning job insurance and income because workers are risk-averse with respect to income. Workers dislike the risk arising from income and fluctuations of employment more than the firms. Consequently, firms offer workers an implicit contract that is partially an income and job insurance contract and partially an employment contract. According to Baily and Azariades, such contracts lead to rigidity in real wages that are not affected by fluctuations in business conditions and employment levels during a recession.

(C) Insider and Outsider Theory

The insider and outsider theory of labour market was developed by A. Lindback and D. Snower.[2] This theory assumes that there are frictions and imperfections in the labour market that act to divide it in terms of employment opportunities.

Insiders are those workers who already have jobs and outsiders are those who are unemployed in the labour market. Insiders are represented by unions who have more say in wage bargaining than the outsiders. Unions negotiate the real wage with firms and set it higher than the market-clearing level so that the outsiders are excluded from jobs leading to involuntary unemployment in the presence of fall in aggregate demand.

Unions use their bargaining power to negotiate wages through turnover costs. Turnover costs relate to the costs of firing, hiring and retaining of new workers. These costs prevent the firms to employ outsiders in place of insiders. Unions can also prevent the entry of outsiders for jobs threatening strikes and work-to-rule. Insiders can also use these costs against outsiders to achieve a higher negotiated wage than the wage at which the outsiders are prepared to work.

---

2. A Lindback and D. Snower, *The Insider and Outsider Theory of Employment and Unemployment*, 1988

However, unions can raise the real wage only up to a certain level because if the real wage is higher than the capacity of the firms to pay, less insiders will be employed, if the aggregate demand falls in the economy.

This theory also explains the persistence of involuntary unemployment if the real wage is set very high above the market-clearing level. This is called *hysteresis*. In times of high involuntary unemployment in a recession, the insiders may use their bargaining power to prevent outsiders form entering into the labour force. Those who become outsiders may lose their influence on wage bargaining contracts because they are no longer union members. Moreover, with persistence involuntary unemployment, it becomes difficult for the long-term unemployed workers to work because their skills have deteriorated. Under the circumstances, a long-period of high involuntary unemployment will tend to become locked-in. This is the hysteresis effect. When outsiders cannot enter the labour market, the hysteresis effect leads to wage stickiness.

(D) Efficiency Wage Theories

In new Keynesian economics, payment of efficiency wages leads to real wage rigidity and the failure of market-clearing mechanism. High wages increase efficiency and productivity of workers. Despite an excess supply of labour, firms do not cut wages even though such a move would increase their profits. Firms also do not cut wages because it would lower productivity and raise costs. So it is in the interest of firms to set the real wage above the market-clearing level. Such a wage is called the efficiency wage.

There are four efficiency wage theories which are explained below.

1. Turnover Costs Theory. According to this theory, the real wage is set to minimise turnover costs of firms. Turnover costs include the costs of firing and hiring workers, and training of new workers. It is profitable for firms of reduce such costs. By paying high real wages above the market-clearing wage, firms can prevent experienced and efficient workers from leaving the firm to join other firms. It can also reduce recruiting costs to replace such workers and costs of training new workers.

2. Selection Theory. Firms do not know the quality of workers at the time of selection. Firms have imperfect information about potential workers at the time of hiring them. Selection procedures being costly, firms always try to select higher quality workers. Good quality workers have a higher reservation (minimum) wage than low quality workers. If a firm pays below the reservation wage, it will not attract good quality workers. By paying a wage higher than the reservation wage, the firm will attract better quality workers. By paying a higher wage, the firm avoids adverse selection (i.e., does not hire low quality workers and reduce productivity of the firm), improves the average quality of workers and increases its productivity. Thus it is advantageous for firms to pay the efficiency wage above the market-clearing wage.

3. On-the-Job Efficiency or Gift Exchange Theory. Another efficiency wage theory is that a real wage above the market-clearing wage improves on-the-job efficiency of workers. A higher wage improves the physical well being of the workers. Their consumption level increases. They can afford better nutritional food and become more healthy. This improves their efficiency and raises their productivity. Moreover, a higher wage acts as an incentive to workers to invest in skill formation so as to move in the upper ranks of the firm. Thus a higher wage is a gift to the workers which they reciprocate with higher efficiency and increased productivity.

4. The Shirking Theory. The shirking theory is based on the premise that firms cannot easily monitor the work effort of workers and that workers themselves decide how hard to work. Two options are available to a worker : shirking and no-shirking. Shirking workers if caught by the management are fired (dismissed). Workers know that if fired, they are not likely to get employment immediately at the existing wage. The firm can encourage workers not to shirk work by paying a higher wage. Thus a wage at which no-shirking occurs is an efficiency wage. Such a wage is an employment cost to workers of being fired. Having a job at a higher wage is advantageous to them than being unemployed. So a higher wage is an inducement to workers not to shirk.

Assumptions

This theory is based on the following assumptions :

1. Workers either have jobs or are unemployed.

2. Those workers who shirk work, are fired.

3. The fired workers do not find employment immediately and remain unemployed for some time.

4. Those workers who receive the efficiency wage do not shirk.

5. There is always unemployment in the labour market at the efficiency wage.

Explanation

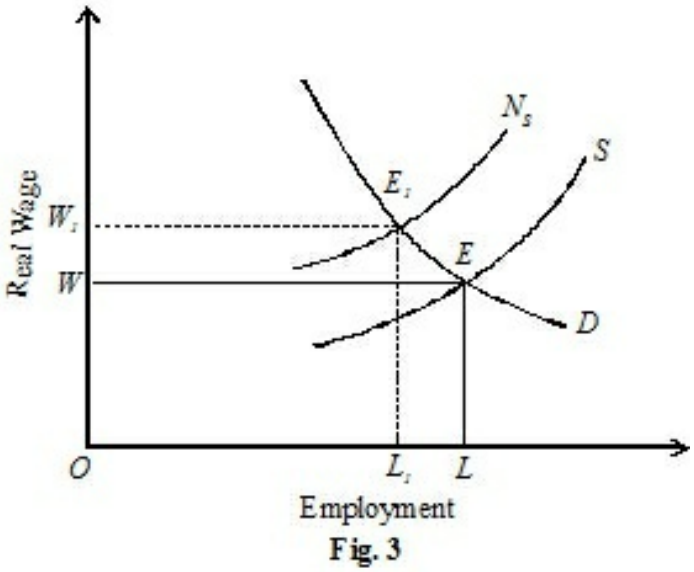


Fig. 3

Given these assumptions, this efficiency wage theory is explained in Figure 3 where *S* is labour supply curve and *D* is the labour demand curve. These curves meet at point *E* where *OL* workers are employed in the firm who are working at *OW* wage rate. The labour demand curve assumes that these workers do not shirk work. But the management finds that some workers shirk work. So they are fired and are unemployed. Since the firm has incurred cost of recruitment and training of some workers, it will ensure that they do not shirk. For this, it pays a higher wage to them which is the no-shirking or efficiency wage. At this wage, there is always the threat of being fired and becoming unemployed. This induces workers not to shirk. In the figure *Ns* is the no-shirking labour supply curve which intersects the labour demand curve at point E_1. OW_1 is the efficiency wage at which OL_1 workers are employed and LL_1 workers remain unemployed who are fired for shirking. The efficiency wage OW_1 is above the market-clearing wage *OW*.

4. COORDINATION FAILURE

New Keynesian theories of wage and price stickiness have inconsistencies because they neglect constraints and spillovers and focus on single markets, one at a time, in a partial equilibrium framework. Cooper and John[3] show that spillovers and strategic complementarities lead to coordination failure. The term 'strategic complementarities' represents a situation in which the optimal decision of one firm is positively dependent on the decision of the other firm. Strategic complementarities are a necessary condition for certain kinds of multiple equilibria that lead to coordination failure and to aggregate fluctuations.

Coordination failure arises when firms and unions try to fix prices and wages to anticipate the actions of other price and wage setters. If there is a change in nominal demand, no firm will have an incentive to change its price exactly in the same proportion unless it believes that other firms will do so immediately. Similarly, unions bargaining for wages will be concerned about the wages other unions may negotiate. But such settings of prices and wages are not possible and they lead to coordination failure.

Suppose there are two firms A and B producing potentially related goods whose demand falls. Each firm has to decide how much price it should cut so as to maximise its profit. The decision to cut its price and profit more would depend on the decision made by the other firm. But the price strategy chosen by one firm is not known to the other firm. This is a duopoly game where each firm has the option of earning more profit or less profit by a price cut or no price cut when demand falls and recession starts.

Table 1 shows the strategy followed by each firm in terms of expected profit when it expects the other to make a move about the price. It both firms do not cut their prices in the face of reduced demand, each makes a profit of $20 million and recession starts. If both firms cut their prices, each earns a high profit of $50 million and the recession is averted. But if firm A cuts its price, it earns a low profit of $ 10 million and if firm B does not cut its price, it earns a high profit of $20 million. In this situation, by cutting its price firm A has improved the position of firm B, which can

avoid recession and earn a high profit. This is due to an aggregate demand externality. Now suppose if firm A expects that firm B will cut its price, it will also cut its price and both will earn the highest profit of $50. Both are able to avert the recession. But if each firm *expects* the other to maintain the same price, each will earn $20 million, the recession will continue. All these situations lead to multiple equilibria. However, the last outcome when each firm earns $20 million is due to coordination failure.

Table 1. Pay-off Matrix in $ million

		B's Strategy: Price cut ↓	B's Strategy: No Price cut ↓
A's Strategy	Price cut →	[50] 50	[10] 20
	No → Price cut	[20] 10	[20] 20

Note : Figures in squares relate to Firm A

POLICY IMPLICATIONS OF NEW KEYNESIAN ECONOMICS

New Keynesian economics has the following policy implications:

1. Monetary and Fiscal Policies. The adherents of new Keynesian economics do not hold a unified view of economic policy because of its different strands. They base their models/theories on the real world imperfectly competitive markets with asymmetrical information where wages and prices are sticky. These lead to decrease in demand, economy-wide market failures, recessions and existence of involuntary unemployment.

3. R. Cooper and A. John, "Coordinating Coordination Failures in Keynesian Models," *Q.J.E.*, August, 1988.

The main elements of new Keynesian economics comprising menu costs, staggered contracts, coordination failures and efficiency wages are substantial departures from the market clearing assumption of new

classical economics. So new Keynesian economics provides a rational for government intervention in the form of active monetary and fiscal policies to prevent recessions and fall in output and employment.

When there is decrease in aggregate demand due to sticky wages and contracts, how stabilisation policy can be used to overcome a recession. This is illustrated in Figure 4.

It is assumed that (i) the economy is at full employment level, (ii) unions and firms have rational expectations, and (iii) aggregate demand decreases after wage contracts, based on these expectations, have been signed and before they are renewed.

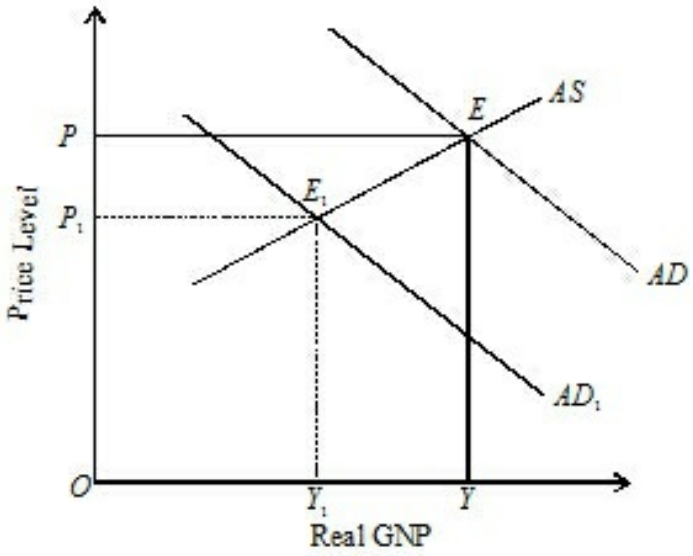


Fig. 4

In the figure, *AD* is the aggregate demand and *AS* is the aggregate supply curve. They meet at point *E* which is the full employment level where *OY* is the real GNP and *OP* is the price level. *A* decrease in aggregate demand due to sticky wages and prices shifts the aggregate demand and curve leftwards to AD_1 which intersects the *AS* curve at E_1. This brings a fall in real GNP to OY_1 and the price to OP_1 leading to a recession. When active monetary and fiscal policies are used, the aggregate demand increases to *AD* and the economy returns to the full employment level *E* with the rise in real GNP to *OY* and price level to *OP*.

Stiglitz favours discretionary monetary and fiscal policies because changing economic circumstances require changes in economic policy. According to him, "If the employment rate becomes high, government must and will do something regardless of what is said."

2. Prices and Incomes Policies. New Keynesian theory also provides for

prices and incomes agreements between firms and unions. In the new Keynesian theory, asymmetries and imperfections in labour market tend to divide the market in terms of employment opportunities. Such imperfections lead to involuntary unemployment. To overcome this problem, prices and incomes policies can increase the power of outsiders in relation to insiders by reducing the effect of market imperfections and limit the use of insiders power thereby reducing unemployment of outsiders.

3. Government and Corporate Policies. Another important implication relates to corporate and government policies when involuntary unemployment persists in the long run after an adverse employment shock. This is called *hysteresis or lagged effects*. In times of high involuntary unemployment in a recession, the insiders may use their bargaining power to prevent outsiders from entering the labour force. Those who become outsiders lose their influence on wage bargaining contracts because they are no longer union members. In such a situation, a long period of high involuntary unemployment will tend to be locked-in. When outsiders cannot enter the labour market, unemployment among them cannot exert pressure on wages which remain sticky.

To reduce hysteresis effects of involuntary unemployment, new Keynesian economists have suggested a number of measures :

(i) Institutional Reforms. Lindback and Snower have suggested a greater role of the government in the bargaining process by centralization of wage bargaining to reduce the power of insiders and to attract outsiders. For this, the government should soften the job security legislation in order to reduce the costs of hiring and firing of workers. It should also reform industrial relations to reduce the frequency of strikes.

(ii) Reform of Benefits System. To provide more employment opportunities to outsiders, the benefits system should be reformed so that the unemployed workers do not depend too much on unemployment insurance, social insurance or 'dole' because such a system encourages workers to shirk jobs. Moreover, efforts should be made to increase labour mobility through efficient employment exchanges and employment bureaus.

(iii) Organisational Changes. Corporations should make organisational changes to increase te power of outsiders by introducing on-the-job training an profit-sharing schemes for workers. By such measures, high involuntary unemployment may persist for a short period.

4. Re-establishment of Policy Effectiveness. New Keynesian theories which emphasise price and wage stickiness, re-establish non-neutrality of money and policy effectiveness. Fisher and Phelps and Taylor have shown that nominal demand disturbances produce real effects in theories that incorporate rational expectations when market-clearing assumption is adandoned. In such theories, monetary policy can help in stabilising the economy. Thus new Keynesian economics re-establishes monetary policy effectiveness as against the policy ineffectiveness of the new classical macroeconomics.

5. Favour Rough or Course Tuning. Monetarists and new classical economists undermined the case for fine-tuning the economy by making frequent changes in monetary and/or fiscal policy to keep GNP at, or near, its full employment level. But new Keynesian economists favoure 'rough' or 'course' tuning where monetary and fiscal policies are used to correct or avoid only large deviations from potential GNP.

6. Existence of Involuntary Unemployment. As against the new classical approach, new Keynesian economists have maintained the existence of involuntary unemployment equilibrium. For example, in efficiency wage models, firms do not cut wages even when there is an excess supply of labour due to persistent unemployment because such a policy would lead to lower efficiency and productivity.

CRITICISMS OF NEW KEYNESIAN ECONOMICS

The new Keynesian economics has been criticised on the following grounds:

1. Lack of a Single Theory/Model. There is no single new Keynesian theory/model as such but a multiplicity of theories/models of price and wage stickiness such as menu-cost approach, asymmnetric information

model, implicit contract theory, insider-outsider model, efficiency wage theories etc. As pointed out by Snowdon and Vane, new Keynesian economics is "an extremely heterogeneous collection of economists and ideas." According to Robert Gordon, "Far from being a set of facts looking for a theory, the new-Keynesian paradigm suffers from too many unrelated theoretical explanations." And Blanchard comments that new Keynesian economics has led to the construction of "too many monsters with few interesting results."

2. Lack of Empirical Research. New Keynesian economics has been criticised for its lack of attention to empirical research. As pointed out by Mankiw, new Keynesian economics has been "successful at the theoretical level but there is a small empirical literature."

3. Models based on Rational Expectations. New Keynesian economics was developed against the new classical assumptions of continuous market clearing and rational expectations. But the new Keynesian models of wage and price stickiness are based on rational expectations. According to Gordon, "Most new-Keynesian models combine rational expectations with maximising behaviour at the level of individual agent. Any attempt to build a model based on irrational behaviour or sub maximising behaviour is viewed as cheating." Greenwald and Stiglitz also point out that "many new Keynesians are not adverse to using the rational expectation assumption when it is convenient to do so."

4. Use of IS-LM Model. Robert King[4] has criticised new Keynesian economists like Mankiw for their faith in the traditional IS-LM model. Because of its treatment of expectations, "the IS-LM model is a hazardous base on which to build positive theories of business cycles and to undertake policy analysis."

Conslusion

Despite these criticisms, new Keynesian economics has shown remarkable resilience of Keynesian economics in facing counter-revolutions against its central doctrines. By emphasising real world imperfectly competitive labour and product markets, and coordination failures, new Keynesian economists have developed macroeconomic models on firm

microeconomic foundations which reject continuous market clearing. But they incorporate rational expectations and natural rate hypotheses in their analysis.

SELECTED READINGS

O.J. Blanchard and S. Fisher, *Lectures on Macroeconomics*, 1989.

G.A. Akerlof and J. Yellen, *Efficiency Wage Models and the Labour Market*, 1986.

R.T. Froyen, *Macroeconomics*, 7/e, 2003

R.J. Gordon, *Macroeconomics*, 6/e, 1993

R.J.Gordon, "What is New-Keynesian *Economics*," *J.E.L.*, September 1990

N.G. Mankiw,*Macroeconomics*, 4/e, 2003

D. Romer, *Advanced Macroeconomics*, 1996.

B. Snowdon and H. Vane," New Keynesian Economics Today," *A.E.*, Spring 1995

QUESTIONS

1. How does new Keynesian economics differ from new classical macroeconomics?
2. How do new Keynesian economists explain nominal price rigidity?
3. Explain the new Keynesian theory of nominal wage stickiness.
4. Critically examine the menu cost hypothesis of sticky nominal prices.
5. Explain how efficiency wage theory provides an explanation of involuntary unemployment as an equilibrium phenomenon.
6. What is coordination failure? Explain how coordination failure arises.

7. What is new Keynesian economics? Discuss its policy implications.
8. "Far from being a set of facts looking for a theory, the new Keynesian paradigm suffers from too many unrelated theoretical explanations." Robert Gordon. In the light of this statement, explain the weaknesses of new Keynesian economics.
9. Write notes on : (a) Staggered wage setting, (b) Insider-Outsider theory, (c) Coordination failure, (d) Hysteresis (e) Implicit contract theory.

---

4. R.G. King, "Will the New Keynesian Macroeconomics Resurrect the IS-LM Model ?" *J.E.P.*, Winter 1993.

PART - IX

MACROECONOMICS IN OPEN ECONOMY

CHAPTER 59

BALANCE OF PAYMENTS: MEANING AND COMPONENTS

MEANING

The balance of payments of a country is a systematic record of all its economic transactions with the outside world in a given year. It is a statistical record of the character and dimensions of the country's economic relationships with the rest of the world. According to Bo Sodersten, "The balance of payments is merely a way of listing receipts and payments in international transactions for a country."[1] B. J. Cohen says, "It shows the country's trading position, changes in its net position as foreign lender or borrower, and changes in its official reserve holding."[2]

STRUCTURE OF BALANCE OF PAYMENTS ACCOUNTS

The balance of payments account of a country is constructed on the principle of double-entry book-keeping. Each transaction is entered on the credit and debit side of the balance sheet. But balance of payments accounting differs from business accounting in one respect. In business accounting, debits (–) are shown on the left side and credits (+) on the right side of the balance sheet. But in balance of payments accounting, the practice is to show credits on the left side and debits on the right side of the balance sheet.

1. Bo Sodersen, *International Economics*, 2/e, 1980.

2. B.J. Cohen, *Balance of Payments Policy*, 1969.

When a payment is received from a foreign country, it is a credit transaction while payment to a foreign country is a debit transaction. The principal items shown on the credit side (+) are exports of goods and services, unrequited (or transfer) receipts in the form of gifts, grants, etc. from foreigners, borrowings from abroad, investments by foreigners in the country, and official *sale* of reserve assets including gold to foreign countries and international agencies. The principal items on the *Debit* side (–) include imports of goods and services, transfer (or unrequited) payments to foreigners as gifts, grants, etc., lending to foreign countries, investments by residents to foreign countries, and official *purchase* of reserve assets or gold from foreign countries and international agencies.

These credit and debit items are shown vertically in the balance of payments account of a country according to the principle of double-entry book-keeping. Horizontally, they are divided into three categories : the current account, the capital account, and the official settlements account or the official reserve assets account.

The balance of payments account of a country is constructed in Table 1.

Table 1. Balance of Payments Account

Credits (+) (*Receipts*)	*Debits* (–) (*Payments*)
1. Current Account	
Exports	*Imports*
(*a*) Goods	(*a*) Goods
(*b*) Services	(*b*) Services
(*c*) Transfer Payments	(*c*) Transfer Payments
2. Capital Account	
(*a*) Borrowings from Foreign Countries	(*a*) Lending to Foreign Countries

(*b*) Direct Investments by Foreign Countries	(*b*) Direct Investments in Foreign Countries

3. Official Settlements Account

(*a*) Increase in Foreign Official Holdings	(*a*) Increase in Official Reserve of Gold and Foreign Currencies

Errors and Omissions

---

1. Current Account. The current account of a country consists of all transactions relating to trade in goods and services and unilateral (or unrequited) transfers. Service transactions include costs of travel and transportation, insurance, income and payments of foreign investments, etc. Transfer payments relate to gifts, foreign aid, pensions, private remittances, charitable donations etc. received from foreign individuals and governments to foreigners.

In the current account, merchandise exports and imports are the most important items. Exports are shown as a positive item and are calculated f.o.b. (free on board) which means that costs of transportation, insurance, etc. are excluded. On the other side, imports are shown as a negative item and are calculated c.i.f. which means that costs, insurance and freight are included. The difference between exports and imports of a country is its *balance of visible trade* or merchandise trade or simply *balance of trade*. If visible exports exceed visible imports, the balance of trade is favourable. In the opposite case when imports exceed exports, it is unfavourable.

It is, however, services and transfer payments or invisible items of the current account that reflect the true picture of the balance of payments account. The balance of exports and imports of services and transfer payments is called the *balance of invisible trade*. The invisible items alongwith the visible items determine the actual current account position. If exports of goods and services exceed imports of goods and services, the balance of payments is said to be *favourable*. In the opposite case, it is unfavourable.

In the current account, the exports of goods and services and the receipts of transfer payments (unrequited receipts) are entered as credits (+) because they represent receipts from foreigners. On the other hand, the imports of goods and services and grant of transfer payments to foreigners are entered as debits (–) because they represent payments to foreigners. The net value of these visible and invisible trade balances is the balance on current account.

2. Capital Account. The capital account of a country consists of its transactions in financial assets in the form of short-term and long-term lendings and borrowings, and private and official investments. In other words, the capital account shows international flow of loans and investments, and represents a change in the country's foreign assets and liabilities. Long-term capital transactions relate to international capital movements with maturity of one year or more and include direct investments like building of a foreign plant, portfolio investment like the purchase of foreign bonds and stocks, and international loans. On the other hand, short-term international capital transactions are for a period ranging between three months and less than one year.

There are two types of transactions in the capital account—private and government. Private transactions include all types of investment : direct, portfolio and short-term. Government transactions consist of loans to and from foreign official agencies.

In the capital account, borrowings from foreign countries and direct investment by foreign countries represent capital inflows. They are positive items or credits because these are receipts from foreigners. On the other hand, lending to foreign countries and direct investments in foreign countries represent capital outflows. They are negative items or debits because they are payments to foreigners. The net value of the balances of short-term and long-term direct and portfolio investments is the *balance on capital account.*

Sodersten and Reed refer to the *external wealth account* of a country which shows the stocks of foreign assets held by the country (positive item) and of domestic assets held by foreign investors (liabilities or negative item). The net value of a country's assets and liabilities is its

balance of indebtedness. If its assets are more than its liabilities, then it is a *net creditor*. If its liabilities are more than its assets, then it is a net debtor.[3]

Basic Balance. The sum of current account and capital account is known as the basic balance.

3. The Official Settlements Account. The official settlements account or official reserve assets account is, in fact, a part of the capital account. But the U.K. and U.S. balance of payments accounts show it as a separate account. "The official settlements account measures the change in nation's liquidity and non-liquid liabilities to foreign official holders and the change in a nation's official reserve assets during the year. The official reserve assets of a country include its gold stock, holdings of its convertible foreign currencies and SDRs, and its net position in the IMF." It shows transactions in a country's net official reserve assets.

---

3. *Bo Sodersten and Geoffery Reed, International Economics,* 3/e, 1994.

Errors and Omissions. Errors and ommissions is a balancing item so that total credits and debits of the three accounts must equal in accordance with the principles of double entry book-keeping so that the balance of payments of a country always balances in the accounting sense.

Is Balance of Payments Always in Equilibrium ?

Balance of payments always balances means that the algebraic sum of the net credit and debit balances of current account, capital account and official settlements account must equal zero. Balance of payments is written as

$$B = R_f - P_f$$

where, B represents balance of payments,

R_f receipts from foreigners,

P_f payments made to foreigners.

When $B = R_f - P_f = 0$, the balance of payments is in equilibrium.

When $R_f - P_f > 0$, it implies receipts from foreigners exceed payments made to foreigners and there is *surplus* in the balance of payments. On the other hand, when $R_f - P_f < 0$ or $R_f < P_f$ there is *deficit* in the balance of payments as the payments made to foreigners exceed receipts from foreigners.

If net foreign lending and investment abroad are taken, a flexible exchange rate creates an excess of exports over imports. The domestic currency depreciates in terms of other currencies. The exports becomes cheaper relatively to imports. It can be shown in equation form :

$$X + B = M + I_f$$

Where X represents exports, M imports, I_f foreign investment, B foreign borrowing

or $$X - M = I_f - B$$

or $$(X - M) - (I_f - B) = 0$$

The equation shows the balance of payments in equilibrium. Any positive balance in its current account is exactly offset by negative balance on its capital account and vice versa. In the accounting sense, the balance of payments always balances. This can be shown with the help of the following equation :

$$C + S + T = C + I + G + (X - M)$$

or $$Y = C + I + G + (X - M) \qquad [\therefore Y = C + S + T]$$

where C represents consumption exepnditure, S domestic saving, T tax receipts, I investment expenditures, G government expenditures, X exports of goods and services, and M imports of goods and services.

In the above equation

$C + S + T$ is GNI or national income (Y), and

$C + I + G = A$,

where A is called 'absorption'.

In the accounting sense, total domestic expenditures ($C + I + G$) must equal current income ($C + S + T$) that is $A = Y$. Moreover, domestic saving (S_d) must equal domestic investment (I_d). Similarly, an export surplus on current account ($X > M$) must be offset by an excess of domestic savings over investment ($S_d > I_d$). Thus the balance of payments always balances in the accounting sense, according to the basic principle of accounting. In the accounting system, the inflow and outflow of a transaction are recorded on the credit and debit sides respectively. Therefore, credit and debit sides always balance. If there is a deficit in the current account, it is offset by a matching surplus in the capital account by borrowings from abroad or/and withdrawing out of its gold and foreign exchange reserves, and vice versa. Thus, the balance of payments always balances in this sense also.

MEASURING DEFICIT OR SURPLUS IN BALANCE OF PAYMENTS

If the balance of payments always balances, then why does a deficit or surplus arise in the balance of payment of a country? It is only when all items in the balance of payments are included that there is no possibility of a deficit or surplus. But if some items are excluded from a country's balance of payments and then a balance is struck, it may show a deficit or surplus.

There are three ways of measuring deficit or surplus in the balance of payments.

First, there is the *basic balance* which includes the current account balance and the long-term capital account balance.

Second, there is the *net liquidity balance* which includes the basic balance and the short-term private non-liquid capital balance, allocation of SDRs, and errors and ommissions.

Third, there is the *official settlements balance* which includes the total net liquid balance and short-term private liquid capital balance.

If the total debits are more than total credits in the current and capital accounts, including errors and omissions, the net *debit balance* measures the *deficit* in the balance of payments of a country. This deficit can be settled with an equal amount of *net credit balance* in the official settlements account. On the contrary, if total credits are more than total debits in the current and capital accounts, including errors and omissions, the *net debit* balance measures the *surplus* in the balance of payments of a country. This surplus can be settled with an equal amount of *net debit* balance in the official settlements account.

The relationship between these balances is summarised in Table 2 below.

TABLE 2

Trade balance	*a*	
Transfer payments balance	*b*	Autonomous
Current Account Balance ..	*c* (= *a* + *b*)	Items
Long-term capital balance	*d*	
Basic Balance	*e* (= *c* + *d*)	
Short-term private non-liquid capital balance	*f*	
Allocation of SDRs	*g*	Accommodating
Errors and omissions	*h*	Items
Net Liquidity Balance.....	*i* (= *e* + *f* + *g* + *h*)	
Short-term private liquid capital balance	*j*	
Official Settlements Balance	*k* (= *i* + *j*)	

Autonomous and Accommodating Items

Each balance would give different figure of the deficit. The items that are included in a particular balance are placed 'above the line' and those excluded are put 'below the line'. Items that are put above the line are called *autonomous items*. Items that are placed below the line are called *settlement or accommodating or compensatory* or *induced items*. All transactions in the current and capital accounts are autonomous items because they are underaken for business or profit motives and are independent of balance of payments considerations. According to Sodersten and Reed, "Transactions are said to be autonomous if their value is determined independently of the balance of payments". Whether there is *BOP* deficit or surplus depends on the balance of autonomous items. If autonomous receipts are less than autonomous payments, *BOP* is in deficit and vice versa.

"Accommodating items on the other hand are determined by the net consequences of the autonomous items", according to Sodersten and Reed. They are in the official reserve account. They are compensating (induced or accommodating) short-term capital transactions which are meant to correct a disequilibrium in the autonomous items of balance of payments.

But it is difficult to determine which item is compensatory and which is autonomous. For instance, in the table given above, the main difference in the three balances is their treatment of short-term capital movements which are responsible for deficit in the balance of payments. The basic balance places short-term private non-liquid capital movements below the line while the net liquid balance puts them above the line. Similarly, the net liquid balance places short-term private liquid capital movements below the line and the official settlements balance puts them above the line. Thus, as pointed out by Sodersten and Reed, "Essentially the distinction between autonomous and accommodating items lies in the motives underlying a transaction, which are almost impossible to determine"[4].

Conclusion. The above analysis is based on the assumption of fixed

exchange rates. Thus a deficit (or surplus in the balance of payments is possible under a system of fixed exchange rates. But under freely floating exchange rates, there can in principle be no deficit (or surplus) in the balance of payments. The country can prevent a deficit or (surplus) by depreciating (or appreciating) its currency. Further, balance of payments always balances in an ex-post accounting sense, according to the basic principle of accounting. Lastly, such a balance of payments can be in equilibrium only if there are no compensating transactions.

---

4. Bo Sodersten and G. Reed, *International Economics*, 3/e, 1994, Italics added.

BALANCE OF TRADE AND BALANCE OF PAYMENTS

The *balance of payments* of a country is a systematic record of its receipts and payments in international transactions in a given year. Each transaction is entered on the credit and debit side of the balance sheet (see Table 1). The principal items on the credit side are : (1) Visible exports which relate to the goods exported for which the country receives payments. (2) Invisible exports which refer to the services rendered by the country to other countries. Such services consist of banking, insurance, shipping, and other services rendered in the form of technical know-how, etc., money spent by tourists and students visiting the country for travel and education, etc. (3) Transfer receipts in the form of gifts received from foreigners. (4) Borrowings from abroad and investments by foreigners in the country. (5) The official sale of reserve assets including gold to foreign countries and international institutions.

The principal items on the debit side are : (1) Visible imports relating to goods imported for which the country makes payments to foreign countries. (2) Invisible imports in the form of payments made by the home country for services rendered by foreign countries. These include all items referred to under (2) in the above para. (3) Transfer payments to foreigners in the form of gifts, etc. (4) Loans to foreign countries, investments by residents in foreign countries, and debt repayments to foreign countries. (5) Official purchase of reserve assets or gold from foreign countries and international institutions.

If the total receipts from foreigners on the credit side exceed the total payments to foreigners on the debit side, the balance of payments is said to be *favourable*. On the other hand, if the total payments to foreigners exceed the total receipts from foreigners, the balance of payments is *unfavourable*.

The *balance of trade* is the difference between the value of goods and services exported and imported. In contains the first two items of the balance of payments account on the credit and the debit side. This is known as "balance of payment on current account." Some writers define the balance of trade as the difference between the value of merchandise exports and imports. Prof. Meade regards this way of defining the balance of trade as wrong and of minor economic significance from the point of view of the national income of the country. In equation form, the balance of payments of $Y = C + I + G + (X - M)$ which includes all transactions which give rise to or exhaust national income. In the equation, Y refers to national income, C to consumption expenditure, I to investment expenditure, G to government expenditure, X to exports of goods and services and M to imports of goods and services. The expression $(X - M)$ denotes the balance of trade. If the difference between X and M is zero, the balance of trade balances. If X is greater than M, the balance of trade is *favourable*, or there is *surplus* balance of trade. On the other hand, if X is less than M, the balance of trade is in *deficit* or is *unfavourable*.

DISEQUILIBRIUM IN BALANCE OF PAYMENTS

A disequilibrium in the *BOP* of a country may be either a deficit or a surplus. A deficit or surplus in *BOP* of a country appears when its autonomous receipts (credits) do not match its autonomous payments (debits). If autonomous credit receipts exceed autonomous debit payments, there is a surplus in the *BOP* and the disequilibrium is said to be *favourable*. On the other hand, if autonomous debit payments exceed autonomous credit receipts, there is a *deficit* in the *BOP* and the disequilibrium is said to be unfavourable or adverse.[5]

Causes of Disequilibrium

There are many factors that may lead to a BOP deficit or surplus:

1. Temporary Changes (or Disequilibrium). There may be a temporary disequilibrium caused by random variations in trade, seasonal fluctuations, the effects of weather on agricultural production, etc. Deficits or surpluses arising from such temporary causes are expected to correct themselves within a short time.

2. Fundamental Disequilibrium. Fundamental disequilibrium refers to a persistent and long-run *BOP* disequilibrium of a country. It is a chronic *BOP* deficit, according to *IMF*. It is caused by such dynamic factors as : (1) Changes in consumer tastes within the country or abroad which reduce the country's exports and increase its imports. (2) Continuous fall in the country's foreign exchange reserves due to supply inelasticities of exports and excessive demand for foreign goods and services. (3) Excessive capital outflows due to massive imports of capital goods, raw materials, essential consumer goods, technology and external indebedness. (4) Low competitive strength in world markets which adversly affects exports. (5) Inflationary pressures within the economy which make exports dearer.

3. Structural Changes (or Disequilibrium). Structural changes bring about disequilibrium in BOP over the long run. They may result from the following factors: (*a*) Technological changes in methods of production of products in domestic industries or in the industries of other countries. They lead to changes in costs, prices and quality of products. (*b*) Import restrictions of all kinds bring about disequilibrium in *BOP*. (*c*) Deficit in *BOP* also arises when a country suffers from deficiency of resources which it is required to import from other countries. (*d*) Disequilibrium in *BOP* may also be caused by changes in the supply or direction of long-term capital flows. More and regular flow of long-term capital may lead to *BOP* surplus, while an irregular and short supply of capital brings *BOP* deficit.

4. Changes in Exchange Rates. Changes in foreign exchange rate in the form of overvaluation or undervaluation of foreign currency lead to *BOP* disequilibrium. When the value of currency is higher in relation to other currencies, it is said to be overvalued. Opposite is the case of an undervalued currency. Overvaluation of the domestic currency makes

foreign goods cheaper and exports dearer in foreign countries. As a result, the country imports more and exports less of goods. There is also outflow of capital. This leads to unfavourable *BOP*. On the contrary, undervaluation of the currency makes *BOP* favourable for the country by encouraging exports and inflow of capital and reducing imports.

5. Cyclical Fluctuations (or Disequilibrium). Cyclical fluctuations in business activity also lead to *BOP* disequilibrium. When there is depression in a country, volumes of both exports and imports fall drastically in relation to other countries. But the fall in exports may be more than that of imports due to decline in domestic production. Therefore, there is an adverse *BOP* situation. On the other hand, when there is boom in a country in relation to other countries, both exports and imports may increase. But there can be either a surplus or deficit in *BOP* situation depending upon whether the country exports more than imports or imports more than exports. In both the cases, there will be disequilibrium in *BOP*.

6. Changes in National Income. Another cause is the change in the country's national income. If the national income of a country increases, it will lead to an increase in imports thereby creating a deficit in its balance of payments, other things remaining the same. If the country is already at full employment level, an increase in income will lead to inflationary rise in prices which may increase its imports and thus bring disequilibrium in the balance of payments.

---

5. The distinction between deficit and disequilibrium in *BOP* should be clearly understood. The former refers to unfavourable or adverse *BOP* as explained above whereas disequilibrium refers to a deficit or surplus in *BOP*.

7. Price Changes. Inflation or deflation is another cause of disequilibrium in the balance of payments. If there is inflation in the country, prices of exports increase. As a result, exports fall. At the same time, the demand for imports increase. Thus increase in export prices leading to decline in exports and rise in imports results in adverse balance of payments.

8. Stage of Economic Development. A country's balance of payments also depends on its stage of economic development. If a country is

developing, it will have a deficit in its balance of payments because it imports raw materials, machinery, capital equipment, and services associated with the development process and exports primary products. The country has to pay more for costly imports and gets less for its cheap exports. This leads to disequilibrium in its balance of payments.

9. Capital Movements. Borrowings and lendings or movements of capital by countries also result in disequilibrium in *BOP*. A country which gives loans and grants on a large scale to other countries has a deficit in its *BOP* on capital account. If it is also importing more, as is the case with the USA, it will have chronic deficit. On the other hand, a developing country brorrowing large funds from other countries and international institutions may have a favourable *BOP*. But such a possibility is remote because these countries usually import huge quanties of food, raw materials, capital goods, etc. and export primary products. Such borrowings simply help in reducing *BOP* deficit.

10. Political Conditions. Political condition of a country is another cause of disequilibrium in *BOP*. Political instability in a country creates uncertainty among foreign investors which leads to the outflow of capital and retards its inflow. This causes disequilibrium in *BOP* of the country. Disequilibrium in *BOP* also occurs in the event of war or fear of war with some other country.

Implications of Disequilibrium

A disequilibrium in the balance of payments whether a deficit or surplus has important implications for a country.

A deficit in the combined current and capital accounts is regarded as undesirable for the country. This is because such a deficit has to be covered by borrowing from abroad or attracting foreign exchange or capital from abroad. This may require paying high interest rates. There is also the danger of withdrawing money by foreigners, as happened in the case of the Asian crisis in the late 1990s. An alternative may be to draw on the reserves of the country which may also lead to a financial crises. Moreover, the reserves of a country being limited, they can be used to pay for *BOP* deficit upto a limit.

But the above analysis of a combined current and capital account deficit is not correct in practice. The reason being that a current account deficit is the same thing as a capital account surplus. However, it is beneficial for a country to have a current account deficit even if it equals capital account surplus in *BOP*. In the short-run, the country may benefit from a higher level of consumption through import of goods and consequently a higher standard of living. But the excess of imports over exports may be financed by foreign investments in the country. These may lead to increased production, employment and income in the country. In the long-run, foreign investors may purchase large assets in the country and thus adversely affect domestic industry as is the case with *MNCs* (multinational corporations).

The current account deficit in *BOP* of a country may have either good or bad effects depending on the nature of an economy.

Take a country where domestic industries are rapidly growing and it has current account *BOP* deficit. These industries offer a high rate of return on their investment. This would, in return, attract foreign investments. As a result, the country would have a capital account surplus due to the inflow of capital and a current account deficit. This current account deficit is good for the economy. No doubt, the external debt of the country increases, but this debt is being utilised to finance the rapid growth of the economy. The real burden of this debt will be very low because it can be repaid out of higher income in the future.

On the contrary, a country having an inefficient and unproductive domestic industry will be adversely affected by its current account *BOP* deficit. The country borrows from abroad to finance the excess of spending over consumption. To attract foreign borrowings, the country will have to pay high interest rates. These will increase the money burden of the debt. The real burden of the debt will also increase because of the low productive capacity of domestic industries. If the current consumption is being financed by foreign borrowings, the wealth of the economy will decline. This, in turn, will lead to either a reduction in domestic expenditure or a change in government policy so as to control the rising debt.

On the other hand, if foreign borrowings are being used to finance real investment, the current account *BOP* deficit will be beneficial for the economy. A higher rate of return on real investment than the interest on foreign borrowings would increase the country's wealth over time through rise in its national income. Thus a current account *BOP* deficit is not always undesirable for a country.

MEASURES TO CORRECT DEFICIT IN BALANCE OF PAYMENTS[6]

When there is a deficit in the balance of payments of a country, adjustment is brought about automatically through price and income changes or by adopting certain policy measures like export promotion, monetary and fiscal policies, devaluation and direct controls. We study these as follows:

1. Adjustment through Exchange Depreciation (Price Effect)

Under flexible exchange rates, the disequilibrium in the balance of payments is automatically solved by the forces of demand and supply for foreign exchange. An exchange rate is the price of a currency which is determined, like any other commodity, by demand and supply. "The exchange rate varies with varying supply and demand conditions, but it is always possible to find an equilibrium exchange rate which clears the foreign exchange market and creates external equilibrium."[7] This is automatically achieved by depreciation of a country's currency in case of deficit in its balance of payments. Depreciation of a currency means that its relative value decreases. Depreciation has the effect of encouraging exports and discouraging imports. When exchange depreciation takes place, foreign prices are translated into domestic prices. Suppose the dollar depreciates in relation to the pound. It means, that the price of dollar falls in relation to the pound in the foreign exchange market. This leads to the lowering of the prices of U.S. exports in Britain and raising of the prices of British imports in the U.S. When import prices are higher in the U.S., the Americans will purchase less goods from the Britishers. On the other hand, lower prices of U.S. exports will increase exports and diminish imports, thereby bringing equilibrium in the balance of payments.

6. If the question relates to disequilibrium in *BOP*, then measures opposite to whatever are given in the paras should also be explained relating to *BOP* surplus, such as appreciation, revaluation, etc. The best answer would be to write at the end of each para : The opposite measures would be adopted in case of *BOP* surplus. Mention the measures in the dotted portion.

2. Devaluation or Expenditure-Switching Policy

Devaluation raises the domestic price of imports and reduces the foreign price of exports of a country devaluing its currency in relation to the currency of another country. Devaluation is referred to as expenditure switching policy because it switches expenditure from imported to domestic goods and services. When a country devalues its currency, the price of foreign currency increases which makes imports dearer and exports cheaper. This causes expenditures to be switched from foreign to domestic goods as the country's exports rise and the country produces more to meet the domestic and foreign demand for goods with reduction in imports. Consequently, the balance of payments deficit is eliminated.

3. Direct Controls

To correct disequilibrium in the balance of payments, government also adopts direct controls which aim at limiting the volume of imports.The government restricts the import of undesirable or unimportant items by levying heavy import duties, fixation of quotas, etc. At the same time, it may allow imports of essential goods duty free or at lower import duties, or fix liberal import quotas for them. For instance, the government may allow free entry of capital goods, but impose heavy import duties on luxuries. Import quotas are also fixed and the importers are required to take licenses from the authorities in order to import certain essential commodities in fixed quantities. In these ways, imports are reduced in order to correct an adverse balance of payments. The government also imposes exchange controls. Exchange controls have a dual purpose. They restrict imports and also control and regulate the foreign exchange. With reduction in imports and control of foreign exchange, visible and invisible imports are reduced. Consequently, an adverse balance of payment is corrected.

4. Adjustment through Capital Movements

A country can use capital imports to correct a deficit in its balance of payments. A deficit can be financed by capital inflows.When capital is perfectly mobile within countries, a small rise in the domestic rate of interest brings a large inflow of capital. The balance of payments is said to be in equilibrium when the domestic interest rate equals the world rate. If the domestic interest rate is higher than the world rate, there will be capital inflows and the balance of payments deficit is corrected.

5. Adjustment through Income Changes

Given the foreign exchange rate and prices in a country, an increase in the value of exports, causes an increase in the incomes of all persons associated with the export industries. These, in turn, create demand for other goods and services within the country. This will raise the incomes of persons engaged in the latter industries and services. This process will continue and the national income increases by the value of the multiplier.

6. Stimulation of Exports and Import Substitutes

A deficit in the balance of payments can also be corrected by encouraging exports. Exports can be encouraged by producing quality products, by increasing exports through increased production and productivity, and by better marketing. They can also be increased by a policy of import substitution. It means that the country produces those goods which it imports. In the beginning, imports are reduced but in the long run exports of such goods start. An increase in exports cause the national income to rise by many times through the operation of the foreign trade multiplier. The foreign trade multiplier expresses the change in income caused by a change in exports. Ultimately, the deficit in the balance of payments is removed when exports rise faster than imports.

---

7. Bo Sodersten, *International Economics*, p. 215.

7. Expenditure-Reducing Policies

A deficit in the balance of payments implies an excess of expenditure over income. To correct it, expenditure and income should be brought into

equality. For this expenditure reducing monetary and fiscal policies are used. A contractionary or tight monetary policy relates to increase in interest rates to reduce money supply and a contractionary fiscal policy relates to reduction in government expenditure and or increase in taxes. Thus expenditure reducing policies reduce aggregate demand through higher taxes and interest rates, thereby reducing expenditure and output. The reduction in expenditure and output, in turn, reduces the domestic price level. This gives rise to switching of expenditure from foreign to domestic goods. Consequently, the country's imports are reduced and the balance of payments deficit is corrected.

EXERCISES

1. Enumerate the principle items in the balance of payments of a country. How can a deficit in the balance of payments be corrected?
2. "Balance of Payments always balances." Elucidate. But how do you explain disequilibrium in balance of payments.
3. Distinguish between balance of payments and balance of trade. How can an unfavourable balance of payments be corrected?
4. What are the causes of an adverse balance of payments? Give suggestions to remove an unfavourable balance of payments?
5. Distinguish between : (*a*) Balance of Current Account and Balance of Capital Account; (*b*) Autonomous Balance and Overall Balance; (*c*) Autonomous Transactions and Accommodating Transactions; (*d*) Deficit and Disequilibrium in Balance of Payments.

CHAPTER 60

ADJUSTMENT MECHANISMS OF BALANCE OF PAYMENTS

INTRODUCTION

When there is a deficit or surplus in *BOP* of a country, it is adjusted through the following mechanisms :

1. Automatic adjustment through price and income changes. Price changes are studied under flexible or floating exchange rates and under the Gold Standard. Income changes are explained in terms of the foreign trade multiplier in the next chapter.

2. Adjustment policies are induced changes to correct disequilibrium in *BOP* by the government of a country. They include expenditure changing and expenditure switching policies, maintaining external and internal balance, direct controls, etc. They are discussed in a separate chapter.

3. There are also approaches to balance of payments which form part of policy measures but are usually discussed separately as Elasticity, Absorption, and Monetary Approaches. They are explained below.

AUTOMATIC PRICE ADJUSTMENT UNDER GOLD STANDARD

Under the international gold standard which operated between 1880-1914, the currency in use was made of gold or was convertible into gold at a fixed rate. The central bank of the country was always ready to buy and

sell gold at the specified price. The rate at which the standard money of the country was convertible into gold was called the *mint price* of gold. This rate was called the *mint parity* or *mint par of exchange* because it was based on the mint price of gold. But the actual rate of exchange could vary above and below the mint parity by the cost of shipping gold between the two countries. To illustrate this, suppose the US had a deficit in its balance of payments with Britain. The difference between the value of imports and exports would have to be paid in gold by US importers because the demand for pounds exceeded the supply of pounds. But the transhipment of gold involved transportation cost and other handling charges, insurance, etc. Suppose the shipping cost of gold from the US to Britain was 3 cents. So the US importers would have to spend $ 6.03 ($ 6 + .03c) for getting £ 1. This could be the exchange rate which was the US *gold export point* or *upper specie point.* No US importer would pay more than $ 6.03 to obtain £ 1 because he could buy $ 6 worth of gold from the US treasury and ship it to Britain at a cost of 3 cents per ounce. Similarly, the exchange rate of the pound could not fall below $ 5.97 in the case of a surplus in the US balance of payments. Thus the exchange rate of $ 5.97to a pound was the US *gold import point* or *lower specie point.* The exchange rate under the gold standard was determined by the forces of demand and supply between the gold points and was prevented from moving outside the gold points by shipments of gold. The main objective was to keep *BOP* in equilibrium. A deficit or surplus in *BOP* under the gold standard was automatically adjusted by the price-specie-flow mechanism. For instance, a *BOP* deficit of a country meant a fall in its foreign exchange reserves due to an outflow of its gold to a surplus country. This reduced the country's money supply thereby bringing a fall in the general price level. This, in turn, would increase its exports and reduce its imports. This adjustment process in *BOP* was supplemented by a rise in interest rates as a result of reduction in money supply. This led to the inflow of short-term capital from the surplus country. Thus the inflow of short-term capital from the surplus to the deficit country helped in restoring *BOP* equilibrium.[1]

AUTOMATIC PRICE ADJUSTMENT UNDER FLEXIBLE EXCHANGE RATES (PRICE EFFECT)

Under flexible (or floating) exchange rates, the disequilibrium in the balance of payments is automatically solved by the forces of demand and supply for foreign exchange. An exchange rate is the price of a currency which is determined, like any other commodity, by demand and supply. "The exchange rate varies with varying supply and demand conditions, but it is always possible to find an equilibrium exchange rate which clears the foreign exchange market and creates external equilibrium."[2] This is automatically achieved by a depreciation (or appreciation) of a country's currency in case of a deficit (or surplus) in its balance of payments. Depreciation (or appreciation) of a currency means that its relative value decreases (or increases). Depreciation has the effect of encouraging exports and discouraging imports. When exchange depreciation takes place, foreign prices are translated into domestic prices. Suppose the dollar depreciates in relation to the pound. It means, that the price of dollar falls in relation to the pound in the foreign exchange market. This leads to the lowering of the prices of U.S. exports in Britain and raising of the prices of British imports in the U.S. When import prices are higher in the U.S., the Americans will purchase less goods from the Britishers. On the other hand, lower prices of U.S. exports will increase their sales to Britain. Thus the U.S. exports will increase and imports diminish, thereby bringing equilibrium in the balance of payments.

---

1. For assumptions and criticisms refer to the Mint Parity Theory in the chapter "Foreign Exchange Rate.'

2. Bo Sodersten, *International Economics*, p. 215.

Its Assumptions

This analysis is based on the following assumptions :

1. There are two countries Britain and U.S.

2. Both are on flexible exchange rate system.

3. *BOP* disequilibrium is automatically adjusted by changes in exchange rates.

4. Prices are flexible in both the countries.

5. There is free trade between the two countries.

Explanation

Given these assumptions, the adjustment process is explained in terms of Figure 1 where *D* is the U.S. demand curve of foreign exchange representing its demand for British imports, and *S* is the U.S. supply curve of foreign exchange representing its exports to Britain. At *P* the demand and supply of the U.S. foreign exchange is in equilibrium where the rate of exchange between U.S. dollar and British pound is *OE* and the quantity of exchange is *OQ*.

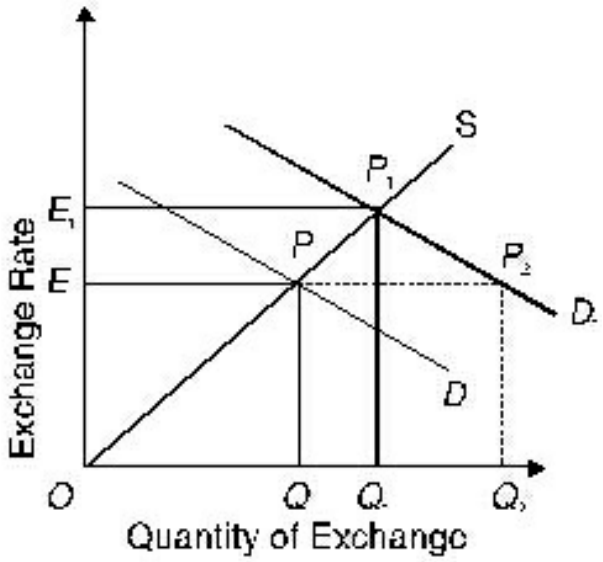


FIG. 1

Suppose disequilibrium develops in the balance of payments of the U.S. in relation to Britain. This is shown by a shift in the demand curve from *D* to D_1 and the incipient deficit equals PP_2. This means an increase in the U.S. demand for British imports which leads to an increase in the demand for pound. This implies depreciation of the U.S. dollar and appreciation of the British pound. As a result, import prices of British goods rise in the U.S. and the prices of U.S. exports fall. This tends to bring about a new equilibrium at P_1 and a new exchange rate at OE_1 whereby the deficit in the balance of payments is eliminated. The demand for foreign exchange equals the supply of foreign exchange at OQ_1 and the balance of payments is in equilibrium.

When the exchange rate rises to OE_1, U.S. goods become cheaper in Britain and British goods become expensive in U.S. in terms of dollar. As a result of changes in *relative prices*, the lower prices of U.S. goods

increase the demand for them in Britain, shown by the new demand curve D_1 . This tends to raise the U.S. exports to Britain which is shown as the movement from *P* to P_1 along the supply curve *S*. At the same time, the higher price of British goods in terms of dollars tends to reduce demand for British goods and to switch demand to domestic goods in the U.S. This leads to the movement from P_2 to P_1 along the new demand curve D_1. Thus the *incipient* deficit PP_2 in *BOP* is removed by increase in the foreign exchange supplied by QQ_1 and decrease in the foreign exchange demanded by Q_2Q_1 so that *BOP* equilibrium is achieved at the exchange rate OE_1 whereby OQ_1 foreign exchange is supplied and demanded.

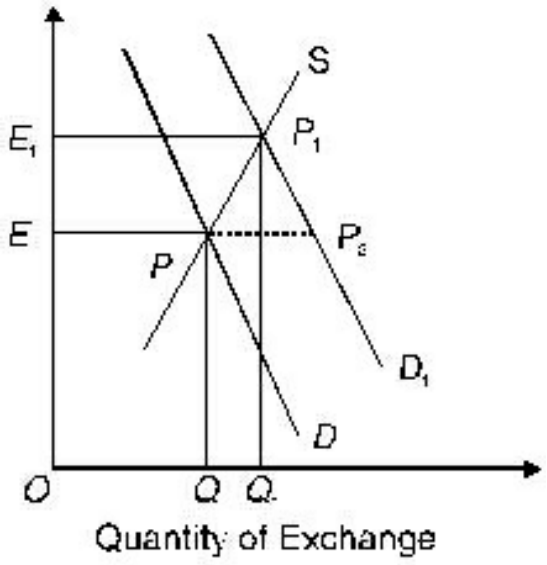


FIG. 2

The above analysis is based on the assumption of *relative elasticities* of demand and supply of foreign exchange. However, in order to measure the full effect of depreciation on relative prices in the two countries, it is not sufficient for demand and supply conditions to be relatively elastic. What is important is *low elasticities* of demand and supply of foreign exchange. This is illustrated in Figure 2 where the original *less elastic* demand and supply curves of foreign exchange are *D* and *S* respectively which intersect at *P* and the equilibrium exchange rate is *OE*. Now a deficit in the balance of payments develops equals to PP_2. Since the *elasticities* of demand and supply of foreign exchange are very low (inelastic), it requires a very large amount of depreciation of the dollar and the appreciation of the pound for the restoration of the equilibrium. The equilibrium will be established through relative price movements in the two countries, as explained above, at P_1 with a very high rate of foreign exchange OE_1. But such a high rate of depreciation would lead to very high price changes in the two countries thereby tending to disrupt their economies.

Its Criticism. The practical use of flexible exchange rates is severely limited. *Depreciation* and *appreciation* lead to fall and rise in prices in the countries adopting them. They lead to severe depressions and inflations respectively. Further, they create insecurity and uncertainty. This is more due to speculation in foreign exchange which destabilises the economies of countries adopting flexible exchange rates. Governments, therefore, favour fixed exchange rates which require adjustments in the balance of payments by adopting policy measures.

THE ELASTICITY APPROACH

Marshall-Lerner Condition

The elasticity approach to *BOP* is associated with the Marshall-Lerner condition which was worked out independently by these two economists. It studies the conditions under which exchange rate changes restore equilibrium in *BOP* by *devaluing* a country's currency. This approach is related to the price effect of devaluation.

Assumptions

This analysis is based on the following assumptions :

1. Supplies of exports are perfectly elastic.

2. Product prices are fixed in domestic currency.

3. Income levels are fixed in the devaluing country.

4. The supply of imparts are large.

5. The price elasticities of demand for exports and imports are *arc* elasticities.

6. Price elasticities refer to absolute values.

7. The country's current account balance equals its trade balance.

The Theory

Given these assumptions, when a country devalues its currency, the domestic prices of its imports are raised and the foreign prices of its exports are reduced. Thus devaluation helps to improve *BOP* deficit of a country by increasing its exports and reducing its imports. But the extent to which it will succeed depends on the country's price elasticities of domestic demand for imports and foreign demand for exports. This is what the Marshall-Lerner condition states: *when the sum of price elasticities of demand for exports and imports in absolute terms is greater than unity, devaluation will improve the country's balance of payments, i.e.*

$$e_x + e_m > 1$$

where e_x is the demand elasticity of exports and E_m is the demand elasticity for imports. On the contrary, if the sum of price elasticities of demand for exports and imports, in absolute terms, is less unity, $e_x + e_m < 1$, devaluation will *worsen* (increase the deficit) the *BOP*. If the sum of these elasticities in absolute terms is equal to unity, $e_x + e_m = 1$, devaluation has no effect on the *BOP* situation which will remain unchanged.[3]

The following is the *process* through which the Marshall-Lerner condition operates in removing *BOP* deficit of a devaluing country.

Devaluation reduces the domestic prices of exports in terms of the foreign currency. With low prices, exports increase. The extent to which they increase depends on the demand elasticity for exports. It also depends on the nature of goods exported and the market conditions. If the country is the sole supplier and exports raw materials or perishable goods, the demand elasticity for its exports will be low. If it exports machinery, tools and industrial products in competition with other countries, the elasticity of demand for its products will be high, and devaluation will be successful in correcting a deficit.

Devaluation has also the effect of increasing the domestic price of imports

which will reduce the import of goods. By how much the volume of imports will decline depends on the demand elasticity of imports. The demand elasticity of imports, in turn, depends on the nature of goods imported by the devaluing country. If it imports consumer goods, raw materials and inputs for industries, its elasticity of demand for imports will be low. It is only when the import elasticity of demand for products is high that devaluation will help in correcting a deficit in the balance of payments.

Thus it is only when the sum of the elasticity of demand for exports and the elasticity of demand for imports is greater than one that devaluation will improve the balance of payments of a country devaluing its currency.

The J-Curve Effect. Empirical evidence shows that the Marshall-Lerner condition is satisfied in the majority of advanced countries. But there is a general consensus among economists that both demand-supply elasticities will be greater in the long run than in the short run. The effects of devaluation on domestic prices and demand for exports and imports will take time for consumers and producers to adjust themselves to the new situation. The short-run price elasticities of demand for exports and imports are lower and they do not satisfy the Marshall-Lerner condition. Therefore, to begin with, devaluation makes the *BOP* worse in the short-run and then improves it in the long-run. This traces a J-shaped curve through time. This is known as the *J-curve effect* of devaluation. This is illustrated in Fig. 3 where time is taken on the horizontal axis and deficit-surplus on the vertical axis. Suppose devaluation takes place at time *T*. In the beginning, the curve *J* has a big loop which shows increase in *BOP* deficit beyond *D*. It is only after time T_1 that it starts sloping upwards and the deficit begins to reduce. At time T_2 there is equilibrium in *BOP* and then the surplus arises from T_2 to *J*. If the Marshall-Lerner condition is not satisfied, in the long run the J-curve will flatten out to *F* from T_2.

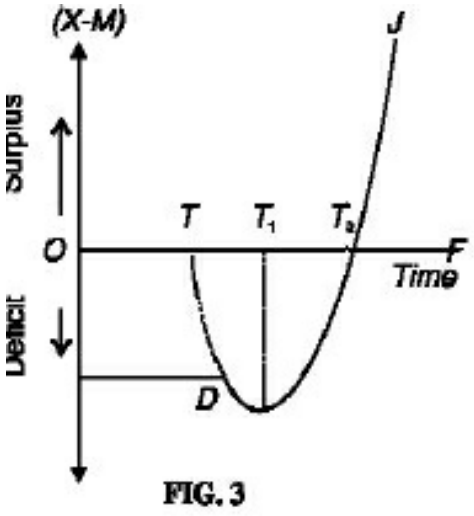


FIG. 3

3. For the derivation of the Marshall-Lerner condition interested students may refer to C.P. Kindleberger, *op. cit.*, Appendix H.

However, in case the country is on a flexible exchange rate, *BOP* will get worse when there is devaluation of its currency. Due to devaluation, there is excess supply of currency in the foreign exchange market which may go on depreciating the currency. Thus the foreign exchange market becomes unstable and the exchange rate may overshoot its long-run value.

Its Criticisms

The elasticity approach based on the Marshall-Lerner condition has the following defects.

1. Misleading. The elasticity approach which applies the Marshallian concept of elasticity to solve *BOP* deficit is misleading. This is because it has relevance only to incremental change along a demand or supply curve and to problems dealing with shifts in these curves. Moreover, it assumes constant purchasing power of money which is not relevant to devaluation of the country's currency.

2. Partial Elasticities. The elasticity approach has been criticised by Alexander because it uses partial elasticities which exclude all factors except relative prices and quantities of exports and imports. This is applicable only to single-commodity trade rather than to a multi-commodity trade. It makes this approach unrealistic.

3. Supplies not Perfectly Elastic. The Marshall-Lerner condition assumes perfectly elastic supplies of exports and imports. But this assumption is unrealistic because the country may not be in a position to increase the supply of its exports when they become cheap with devaluation of its currency.

4. Partial Equilibrium Analysis. The elasticity approach assumes domestic price and income levels to be stable within the devaluing country. It, further, assumes that there are no restrictions in using additional resources into production for exports. These assumptions show that this analysis is based on the partial equilibrium analysis. It, therefore,

ignores the feedback effects of a price change in one product on incomes, and consequently on the demand for goods. This is a serious defect of the elasticity approach because the effects of devaluation always spread to the entire economy.

5. Inflationary. Devaluation can lead to inflation in the economy. Even if it succeeds in improving the balance of payments, it is likely to increase domestic incomes in export and import-competing industries. But these increased incomes will affect the *BOP* directly by increasing the demand for imports, and indirectly by increasing the overall demand and thus raising the prices within the country.

6. Ignores Income Distribution. The elasticity approach ignores the effects of devaluation on income distribution. Devaluation leads to the reallocation of resources. It takes away resources from the sector producing non-traded goods to export and import-competing industries sector. This will tend to increase the incomes of the factors of production employed in the latter sector and reduce that of the former sector.

7. Applicable in the Long Run. As discussed above in the J-curve effect of devaluation, the Marshall-Lerner condition is applicable in the long-run and not in the short. This is because it takes time for consumers and producers to adjust themselves when there is devaluation of the domestic currency.

8. Ignores Capital Flows. This approach is applicable to *BOP* on current account or balance of trade. But *BOP* deficit of a country is mainly the result of the outflow of capital. It thus ignores *BOP* on capital account. Devaluation as a remedy is meant to cut imports and the outflow of capital and increase exports and the inflow of capital.

Conclusion. There has been much controversy over the Marshall-Lerner condition for improvements in the balance of payments. Economists tried to measure demand elasticities in international trade. Some economists found low demand elasticities and others high demand elasticities. Accordingly, the former suggested that devaluation was not an effective method while the latter suggested that it was a potent mechanism of balance of payments adjustment. But it is difficult to generalise due to

these diverse findings on account of differences in the volume and structure of foreign trade.

THE ABSORPTION APPROACH

The absorption approach to balance of payments is general equilibrium in nature and is based on the Keynesian national income relationships. It is, therefore, also known as the Keynesian approach. It runs through the *income effect of devaluation* as against the price effect to the elasticity approach. The theory states that if a country has a deficit in its balance of payments, it means that people are 'absorbing' more than they produce. Domestic expenditure on consumption and investment is greater than national income. If they have a surplus in the balance of payments, they are absorbing less. Expenditure on consumption and investment is less than national income. Here the *BOP* is defined as the difference between national income and domestic expenditure.

This approach was developed by Sydney Alexander.[4] The analysis can be explained in the following form

$$Y = C + I_d + G + X - M \qquad ...(1)$$

where Y is national income, C is consumption expenditure, I_d total domestic investment, G is autonomous government expenditure, X represents exports and M imports.

The sum of $(C + I_d + G)$ is the total absorption designated as A, and the balance of payments $(X - M)$ is designated as B. Thus Equation (1) becomes

$$Y = A + B$$

or

$$B = Y - A \qquad ...(2)$$

which means that *BOP* on current account is the difference between national income (Y) and total absorption (A). *BOP* can be improved by

either increasing domestic income or reducing the absorption. For this purpose, Alexander advocates devaluation because it acts both ways. *First*, devaluation increases exports and reduces imports, thereby increasing the national income. The additional income so generated will further increase income via the multiplier effect. This will lead to an increase in domestic consumption. Thus the net effect of the increase in national income on the balance of payments is the difference between the total increase in income and the induced increase in absorption, i.e.,

---

4. S. Alexander "Effects of Devaluation on a Trade Balance," *I.M.F. Staff Papers*, April 1952. But it was R.F. Harrod who was the first economist to apply the Keynesian analysis to *BOP* in his *International Trade*.

$$\Delta B = \Delta Y - \Delta A \qquad ...(3)$$

Total absorption (ΔA) depends on the marginal propensity to absorb[5] when there is devaluation. This is expressed as a. Devaluation also directly affects absorption through the change in income which we write as D. Thus

$$\Delta A = a\Delta Y + \Delta D \qquad ...(4)$$

Substituting equation (4) in (3), we get

$$\Delta B = \Delta Y - a\Delta Y - \Delta D$$

or

$$\Delta B = (1 - a)\,\Delta Y - \Delta D \qquad ...(5)$$

The equation points toward three factors which explain the effects of devaluation on *BOP*. They are : (*i*) the marginal propensity to absorb (a), (*ii*) change in income (ΔY), and (*iii*) change in direct absorption (ΔD). It may be noted that since a is the marginal propensity (*MP*) to absorb, $(1 - a)$ is the propensity to hoard or save. These factors, in turn, are influenced by the existence of unemployed or idle resources and fully employed resources in the devaluing country.

Effects of Devaluation on BOP

1. MP to Absorb. To take the *MP* to absorb, it it is less than unity ($a < 1$), with idle resources in the country, devaluation will increase exports and reduce imports. Output and income will rise and *BOP* on current account will improve. If, on the other hand, $a > 1$, there will be an adverse effect of devaluation on *BOP*. It means that people are absorbing more or spending more on consumption and investing more. In other words, they are spending more than the country is producing. In such a situation, devaluation will not increase exports and reduce imports, and *BOP* situation will worsen.

Under conditions of full employment if $a > 1$, the government will have to follow expenditure reducing policy measures along with devaluation whereby the resources of the economy are so reallocated as to increase exports and reduce imports. Ultimately, *BOP* situation will improve.

2. Income Effects. Let us take the income effects of devaluation. If there are idle resources, devaluation increases exports and reduces imports of the devaluing country. With the expansion of export and import-competing industries, income increases. The additional income so generated in the economy will further increase income via the multiplier effect. This will lead to improvement in *BOP* situation. If resources are fully employed in the economy, devaluation cannot correct an adverse *BOP* because national income cannot rise. Rather, prices may increase thereby reducing exports and increasing imports, thereby worsening the *BOP* situation.

3. Terms of Trade Effect. The effect of devaluation on national income is also through its effects on the terms of trade. The conditions under which devaluation *worsens* the terms of trade, national income will be *adversely* affected, and *vice versa*. Generally, devaluation worsens the terms of trade because the devaluing country has to export more goods in order to import the same quantity as before. Consequently, the trade balance deteriorates and national income declines. If prices are fixed in buyer's (other country's) currency after devaluation, the terms of trade improve because exports increase and imports decline. The importing country pays more for increased exports of the devaluing country than it

receives from its imports. Thus the trade balance of the devaluing country improves and its national income rises.

---

5. The marginal propensity to absorb (*a*) is the sum of the marginal propensity to consume and marginal propensity to invest. *G* has no role as it is autonomous government expenditure.

4. Direct Absorption. Devaluation affects direct absorption in a number of ways. It the devaluing country has idle resources, an expansionary process will start with exports increasing and imports declining. Consequently, income will rise and so will absorption. If the increase in absorption in less than the rise in income, *BOP* will improve. Generally, the effect of devaluation on direct absorption is not significant in a country with idle resources.

If the economy is fully employed and has also a *BOP* deficit, national income cannot be increased by devaluing the currency. So an improvement in *BOP* can be brought about by reduction in direct absorption. Domestic absorption can fall automatically as a result of devaluation due to real cash balance effect, money illusion and income redistribution.

5. Real Cash Balance Effect. When a country devalues its currency, its domestic prices rise. If the money supply remains constant, the real value of cash balances held by the people falls. To replenish their cash balances, people start saving more. This can be possible only by reducing their expenditure or absorption. This is the real cash balance effect of devaluation.

If people hold assets and when devaluation reduces their real cash balances, they sell them. This reduces the prices of assets and increases the interest rate. This, in turn, will reduce investment and consumption, given the constant money supply. As a result, absorption will be reduced. This is the *asset effect* of real cash balance effect of devaluation.

6. Money Illusion Effect. The presence of money illusion also tends to reduce direct absorption. When prices rise due to devaluation, consumers think their real incomes have fallen, even though their money incomes have risen. They have the money illusion under whose influence they

reduce their consumption expenditure or direct absorption.

7. Income Re-distribution Effect. Direct absorption falls automatically if devaluation redistributes income in favour of people with high marginal propensity to save and against thosee with high marginal propensity to consume. If the marginal propensity to consume of workers is higher than those of profit-earners, absorption will be reduced. Further, when money incomes of lower income groups increase with devaluation, they enter the income tax bracket. When they start paying income tax, they reduce their consumption as compared with higher income groups which are already paying the tax. This leads to reduction in absorption in case of the former.

Income redistribution also takes place between *production sectors* after devaluation. Those sectors whose prices rise more than their costs of production earn more profits than the other sectors whose costs rise more than their prices. Thus the effect of devaluation will be to redistribute income in favour of the former sectors.

Devaluation will also redistribute income in favour of sectors producing and selling *traded goods* and against *non-traded goods* sectors. Prices of traded goods rise more than that of non-traded goods. As a result, profits of producers and traders and wages of workers producing traded goods rise more as compared to those engaged in non-traded goods.

8. Expenditure-Reducing Policies. Direct absorption is also reduced if the government adopts expenditure-reducing monetary-fiscal policies which are deflationary. They will make devaluation successful in reducing *BOP* deficit. But they will create unemployment in the country.

Its Criticisms

The absorption approach to *BOP* deficit has been criticised on the following grounds :

1. Neglects Price Effects. This approach neglects the price effects of devaluation which are very important.

2. Calculation Difficult. Analytically, it appears to be superior to the

elasticity approach but propensities to consume, save and invest cannot be accurately calculated.

3. Ignores Effects on Other Countries. The absorption approach is weak in that it relies too much on policies designed to influence domestic absorption. It does not study the effects of a devaluation on the absorption of other countries.

4. Not Operative in a Fixed Exchange Rate System. The absorption approach fails as a corrective measure of *BOP* deficit under a fixed exchange rate system. When prices rise with devaluation, people reduce their consumption expenditure. With money supply remaining costant, interest rate rises which brings a fall in output alongwith absorption. Thus devaluation will have little effect on *BOP* deficit.

5. More Emphasis on Consumption. This approach places more emphasis on the level of domestic consumption than on relative prices. A mere reduction in the level of domestic consumption for reducing absorption does not mean that resources so released will be redirected for improving *BOP* deficit.

THE MONETARY APPROACH

The monetary approach to the balance of payments is an explanation of the overall balance of payments. It explains changes in balance of payments in terms of the demand for and supply of money. According to this approach, “a balance of payments deficit is always and everywhere a monetary phenomenon.” Therefore, it can only be corrected by monetary measures.

Its Assumptions

This approach is based on the following assumptions:

1. The ‘law of one price’ holds for identical goods sold in different countries, after allowing for transport costs.

2. There is perfect substitution in consumption in both the product and

capital markets which ensures one price for each commodity and a single interest rate across countries.

3. The level of output of a country is assumed exogenously.

4. All countries are assumed to be fully employed where wage price flexibility fixes output at full employment.

5. It is assumed that under fixed exchange rates the sterilisation of currency flows is not possible on account of the law of one price globally.

6. The demand for money is a stock demand and is a stable function of income, prices, wealth and interest rate.

7. The supply of money is a multiple of monetary base which includes domestic credit and the country's foreign exchange reserves.

8. The demand for nominal money balances is a positive function of nominal income.

The Theory

Given these assumptions, the monetary approach can be expressed in the form of the following relationship between the demand for and supply of money :

The demand for money (M_D) is a stable function of income (Y), prices (P) and rate of interest (i)

$$M_D = f\,(Y, P, i) \qquad ...(1)$$

The money supply (M_S) is a multiple of monetary base (m) which consists of domestic money (credit) (D) and country's foreign exchange reserves (R). Ignoring m for simplicity which is a constant,

$$M_S = D + R \qquad ...(2)$$

Since in equilibrium the demand for money equals the money supply,

$$M_D = M_S \qquad ...(3)$$

or $$M_D = D + R \qquad [\, M_S = D + R] \qquad ...(4)$$

A balance of payments deficit or surplus is represented by changes in the country's foreign exchange reserves. Thus

$$\Delta R = \Delta M_D - \Delta D \qquad ...(5)$$

or $$\Delta R = B \qquad ...(6)$$

where B represents balance of payments which is equal to the difference between change in the demand for money (ΔM_D) and change in domestic credit (ΔD).

A balance of payments deficit means a *negative B* which reduces R and the money supply. On the other hand, a surplus means a positive B which increases R and the money supply. When $B = O$, it means *BOP* equilibrium or no disequilibrium of *BOP*.

The automatic adjustment mechanism in the monetary approaches is explained under both the fixed and flexible exchange rate systems.

Under the *fixed exchange rate system*, assume that $M_D = M_S$ so that *BOP* (or B) is zero. Now suppose the monetary authority increases domestic money supply, with no change in the demand for money. As a result, $M_S > M_D$ and there is a *BOP* deficit. People who have larger cash balances increase their purchases to buy more foreign goods and securities. This tends to raise their prices and increase imports of goods and foreign assets. This leads to increase in expenditure on both current and capital accounts in *BOP*, thereby creating a *BOP* deficit. To maintain a fixed exchange rate, the monetary authority will have to sell foreign exchange reserves and buy domestic currency. Thus the outflow of foreign exchange reserves

means a fall in R and in domestic money supply. This process will continue until $M_S = M_D$ and there will again be *BOP* equilibrium.

On the other hand, if $M_S < M_D$ at the given exchange rate, there will be a *BOP* surplus. Consequently, people acquire the domestic currency by selling goods and securities to foreigners. They will also seek to acquire additional money balances by restricting their expenditure relatively to their income. The monetary authority on its part, will buy excess foreign currency in exchange for domestic currency. There will be inflow of foreign exchange reserves and increase in domestic money supply. This process will continue until $M_S = M_D$ and *BOP* equilibrium will again be restored. Thus a *BOP* deficit or surplus is a temporary phenomenon and is self-correcting (or automatic) in the long-run.

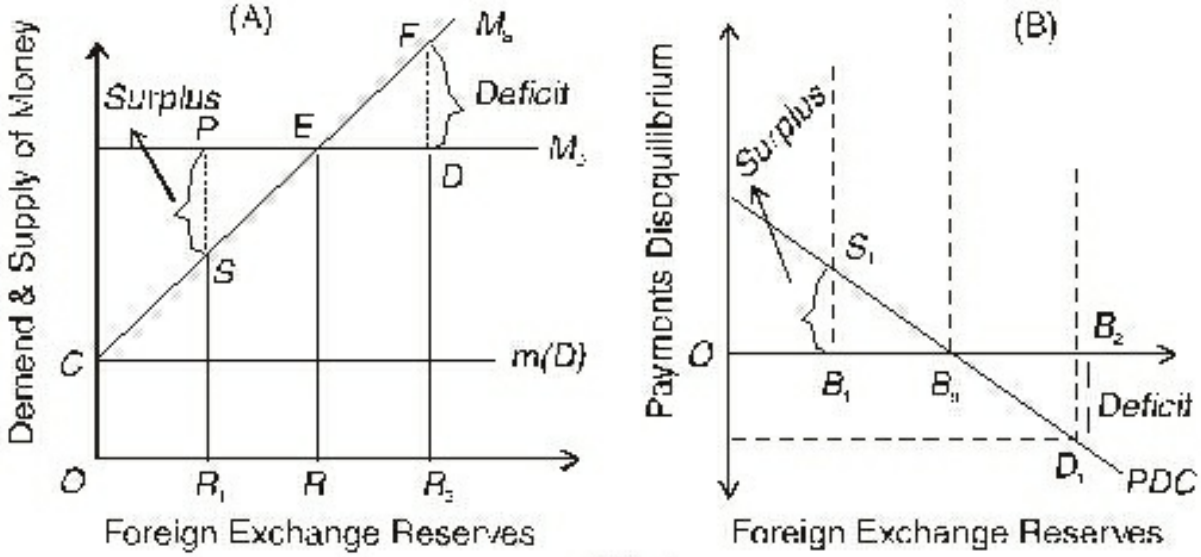


FIG. 4

This is explained in Fig. 4 In Panel *(A)* of the figure, M_D is the stable money demand curve and M_S is the money supply curve. The horizontal line *m* (*D*) represents the monetary base which is a multiple of domestic credit, *D* which is also constant. This is the domestic component of money supply that is why the M_S curve starts from point *C*.

M_S and M_D curves intersect at point *E* where the country's balance of payments is in equilibrium and its foreign exchange reserves are *OR*. In

Panel (*B*) of the figure, *PDC* is the payments disequilibrium curve which is drawn as the vertical difference between M_S and M_D curves of Panel (*A*). As such, point B_0 in Panel (*B*) corresponds to point *E* in Panel (*A*) where there is no disequilibrium of balance of payments.

If $M_S < M_D$ there is *BOP* surplus of *SP* in Panel (*A*). It leads to the inflow of foreign exchange reserves which rise from OR_1 to *OR* and increase the money supply so as to bring *BOP* equilibrium at point *E*. On the other hand, if $M_S > M_D$, there is deficit in *BOP* equal to *DF*. There is outflow of foreign exchange reserves which decline from OR_2 to *OR* and reduce the money supply so as to reestablish *BOP* equilibrium at point *E*. The same process is illustrated in Panel (*B*) of the figure where *BOP* disequilibrium is self-correcting or automatic when B_1S_1 surplus and B_2D_1 deficit are equal.

Under a system of *flexible* (or *floating*) *exchange rates*, when $B = O$, there is no change in foreign exchange reserves (*R*). But when there is a *BOP* deficit or surplus, changes in the demand for money and exchange rate play a major role in the adjustment process without any inflow or outflow of foreign exchange reserves. Suppose the monetary authority increases the money supply ($M_S > M_D$) and there is a *BOP* deficit. People having additional cash balances buy more goods thereby raising prices of domestic and imported goods. There is depreciation of the domestic currency and a rise in the exchange rate. The rise in prices, in turn, increases the demand for money thereby bringing the equality of M_D and M_S without any outflow of foriegn exchange reserves. The opposite will happen when $M_D > M_S$, there is fall in prices and appreciation of the domestic currency which automatically eliminates the excess demand for money. The exchange rate falls until $M_D = M_S$ and *BOP* is in equilibrium without any inflow of foreign exchange reserves.

Its Criticisms

The monetary approach to the balance of payments has been criticised on a number of counts:

1. Demand for Money not Stable. Critics do not agree with the assumption of stable demand for money. The demand for money is stable in the long run but not in the short run when it shows less stability.

2. Full Employment not Possible. Similarly, the assumption of full employment is not acceptable because there exists involuntary unemployment in countries.

3. One Price Law Invalid. Frankel and Johnson[6] are of the view that the law of one price holds for identical goods sold is invalid. This is because when factors of production are drawn into sectors producing non-trading goods, the excess demand for non-traded goods will spill over into reduced supplies of traded goods. This will lead to higher imports, and disturb the law of one price for all traded goods.

4. Market Imperfections. There are also market imperfections which prevent the law of one price from working properly in many markets for traded goods. There may be price differentials due to the lack of information about overseas prices and trade regulations faced by traders.

5. Sterilisation not Possible. The assumption that the sterilisation of currency flows is not possible under fixed exchanged rates, has not been accepted by critics. They argue that "the sterilisation of currency flows is entirely possible if the private sector is willing to adjust the composition of its wealth portfolio with regard to the relative importance of bonds and money balances, or if the public sector is prepared to run a higher budget deficit whenever it has a balance of payments deficit with which to contend."

6. Link between BOP and Money Supply not Valid. The monetary approach is based upon direct link between *BOP* of a country and its total money supply. This has been questioned by economists. The link between the two depends upon the ability of the monetary authority to neutralise the inflows and outflows of foreign exchange reserves when there is *BOP* deficit and surplus. This requires some degree of sterilisation of external flows. But this is not possible due to globalisation of financial markets.

6. L.A. Frankel and H.G. Johnson (eds.), *The Monetary Approach to the Balance of Payments,* 1976.

7. Neglects Short Run. The monetary approach is related to the self correcting long-run equilibrium in *BOP*. This is unrealistic because it fails to describe the short time through which the economy passes to reach the new equilibrium. As pointed out by Prof. Krause, the monetary approach's "concentration on the long-run assumes away all of the problems that make the balance of payments a problem."

8. Neglects Other Factors. This approach neglects all real and structural factors which lead to disequilibrium in *BOP* and concentrates only on domestic credit.

9. Neglects Economic Policy. This approach emphasises the role of domestic credit in bringing *BOP* equilibrium and neglects economic policy measures. According to Prof. Currie, the balance of payments equilibrium can also be "achieved by expenditure-switching policies working through real flows and government budget."

Conclusion. Despite these criticisms, the monetary approach is realistic in that it takes into consideration both domestic money and foreign money. Emphasis is not on relative price changes, but on the extent to which the demand for real money balances will be satisfied from internal sources, through credit creation or from external sources through surplus or deficit in the balance of payments. A balance of payments deficit or surplus can be corrected through changes in money supply and their consequent effects on income and expenditure, or more generally on production and consumption of goods.

EXERCISES

1. Critically discuss the monetary approach to the balance of payments problem.
2. Discuss the alternative automatic adjustment mechanisms for balance of payments. What are their assumptions?
3. Discuss the elasticity approach to the effect of devaluation on balance of payments.

4. Discuss the process of adjustment in the balance of payments through the variations in prices. Explain its limitations.
5. Discuss the role of devaluation in reducing deficits in the balance of payments of a country. State its limitations.
6. Discuss the Marshall-Lerner Condition for effective devaluation.
7. Discuss the mechanism of the absorption approach to the balance of payments adjustment. What are its assumptions and limitations?

CHAPTER 61

BALANCE OF PAYMENTS POLICIES : INTERNAL AND EXTERNAL BALANCE

INTRODUCTION

In this chapter, policy measures are discussed which every government tries to pursue to correct disequilibrium in the balance of payments. They are : (*a*) internal balance which refers to full employment with price stability, and (*b*) the external balance or balance of payments equilibrium. It was Meade who in his *The Balance of Payments*[1] pointed out that to maintain both internal and external balance, a country must control both its aggregate expenditure and the exchange rate. It was, however, Johnson[2] who pointed towards the range of policy instruments for bringing about both internal and external balance. He called them *expenditure changing* and *expenditure switching* policies. Expenditure changing policies are discussed below.

1. J.E. Meade, *The Theory of International Economic Policy,* Vol. I, 1951.

2. H.G. Johnson, *Money, Trade and Economic Growth,* 2/e, 1964.

3. Expenditure Switching Policies are discussed in the last section.

EXPENDITURE CHANGING MONETARY AND FISCAL POLICIES

Expenditure changing policies are intended to change the aggregate expenditure in the economy through appropriate monetary and fiscal policies in order to affect its *BOP* disequilibrium.

To explain these two policies, we use the *IS-LM-BP* technique where the *LM* curve represents monetary policy, the IS curve fiscal policy and the *BP* (balance of payments) curve or *EB* (external balance) curve or *EF* (foreign exchange) curve. *The BP curve traces out those combinations of interest rate and national income that produce BOP equilibrium.* An increase in income will increase imports and increase the *BOP* deficit by increasing the current account deficit. On the other hand, an increase in interest rate will attract foreign investments and there will be capital inflow thereby increasing the capital account surplus.

Any point above and to the left of the BP curve represents surplus in *BOP*, and a point below and the right of the *BP* curve shows a *BOP* deficit. The way expenditure changing monetary and fiscal policies affect *BOP* disequilibrium are discussed below.

Expenditure Changing Monetary Policy

An expenditure changing monetary policy affects the economy through changes in money supply and interest rates. A contractionary monetary policy leads to a *BOP* surplus and an expansionary monetary policy to a *BOP* deficit.

Assumptions. The analysis that follows assumes:

(*a*) There is fixed exchange rate;

(*b*) There is relative capital mobility; and

(*c*) There is no change in government expenditures *i.e.*, the *IS* curve remains unchanged.

Expenditure Reducing Monetary Policy

Suppose there is a *BOP* deficit in the country. This implies an excess of expenditure over income. To correct it, the monetary policy reduces the money supply which increases interest rates thereby reducing investment and output. The reduction in investment (expenditure) and output, in turn, reduces income and aggregate demand for imported goods. There is also a

reduction in the domestic price level which may lead to switching of expenditure from foreign to domestic goods. Consequently, the country's imports are reduced and exports are increased. Thus the current account trade deficit is reduced. Simultaneously, there is reduced outflow of short-term capital with reduction in imports which cuts down the *BOP* deficit. On the other hand, rise in domestic interest rates increase the inflow of capital, thereby completely eliminating the *BOP* deficit. The adjustment process will be just the opposite of the above when the monetary authority adopts an expansionary monetary policy to correct a *BOP* surplus. Expenditure reducing monetary policy and its effects on *BOP* situation are illustrated in Fig. 1.

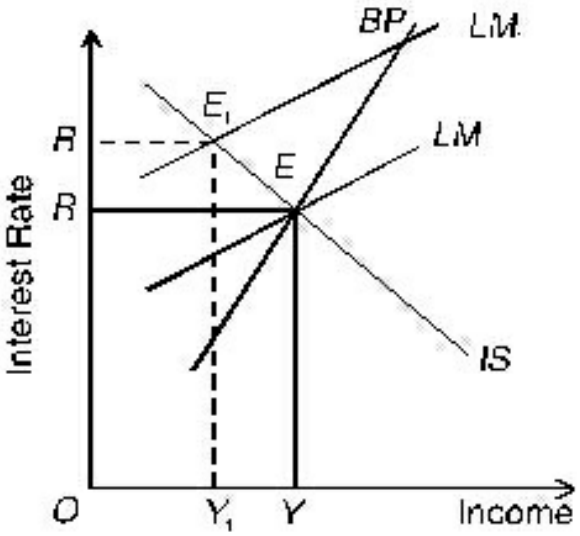


FIG. 1

Given these assumptions, we begin with a situation where the economy is in complete equilibrium with *OR* interest rate and *OY* income level at point *E* of the intersection of *IS-LM-BP* curves. Suppose the domestic money supply is reduced. This will shift the *LM* curve upward to the left to LM_1. The new equilibrium is at E_1. Since E_1 is above and to the left of the *BP* curve, there is *BOP* surplus. There is increase in interest rate from *OR* to OR_1 which generates a capital account surplus with capital inflow. On the other hand, the reduction in income from *OY* to OY_1 will tend to generate a current account surplus because of the reduction in imports. Thus a *contractionary monetary policy leads to a BOP surplus.* However, E_1 does not represent a permanent equilibrium. The *BOP* surplus will increase the domestic money supply and gradually shift the LM_1 curve to the right towards the *LM* curve so that E_1 begins to move down along the *IS* curve to point *E* when the economy is again in *BOP* equilibrium.

Expenditure Increasing Monetary Policy

On the other hand, an *expansionary monetary* policy leads to a *BOP* deficit, as illustrated in Fig. 2. Starting from the complete equilibrium point *E* in the figure, the monetary authority increases the money supply. This will shift the *LM* curve downward to the right to LM_2. The new equilibrium is set at E_2. Since point E_2 is below and to the right of the *BP* curve, there is *BOP* deficit. There is reduction in interest rate from *OR* to OR_2 which generates a capital account deficit with capital outflow. On the other hand, the increase in income from *OY* to OY_2 will tend to generate a current account deficit with increase in imports. Thus an *expansionary monetary policy leads to a BOP deficit.* However, E_2 does not represent a permanent equilibrium. The *BOP* deficit will reduce the domestic money supply and gradually shift the LM_2 curve to the left towards the *LM* curve so that E_2 begins to move up along the *IS* curve to point *E* and the economy is again in *BOP* equilibrium.

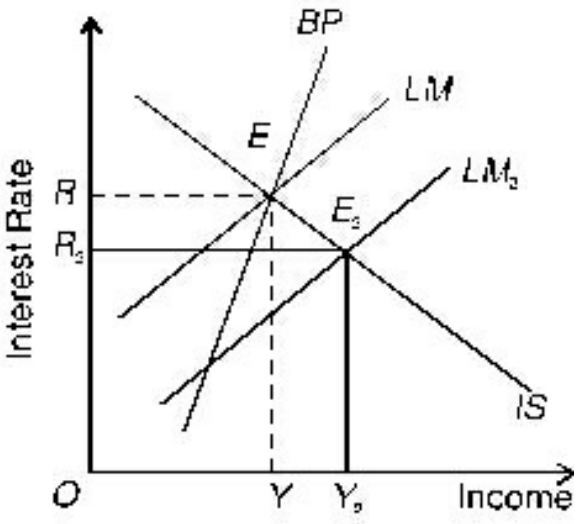


FIG. 2

Expenditure Changing Fiscal Policy

By fiscal policy we mean changing government expenditure or/and taxation. An expansionary fiscal policy tends to increase government expenditure or/and reduce taxes. On the other hand, a contractionary fiscal policy relates to cut in government expenditure or/and increase in taxes.

Expenditure Reducing Fiscal Policy

The effects of an expenditure reducing fiscal policy in correcting a *BOP* deficit are illustrated in Figure 3.

Assumptions. This analysis assumes that :

(*a*) The exchange rate is fixed.

(*b*) There is relative capital mobility.

(*c*) There is no change in monetary policy so that the *LM* curve remains unchanged.

Explanation

Given these assumptions, we start with the equilibrium situation *E* with *OR* interest rate and *OY* income level where *IS-LM-BP* curves intersect in Fig 3. Suppose the government adopts a *contractionary fiscal policy* whereby it reduces its expenditure or/and increases taxes. This shifts the *IS* curve downward to the left to IS_2 which cuts the *LM* curve at point E_2. This point shows fall in interest rate from *OR* to OR_2 which leads to an outflow of capital and to capital account deficit. The income level also falls from *OY* to OY_2 which reduces imports, thereby leading to current account deficit. *Thus the overall effect of a contractionary fiscal policy is to have a BOP deficit* because point E_2 is below and to the right of the *BP* curve.

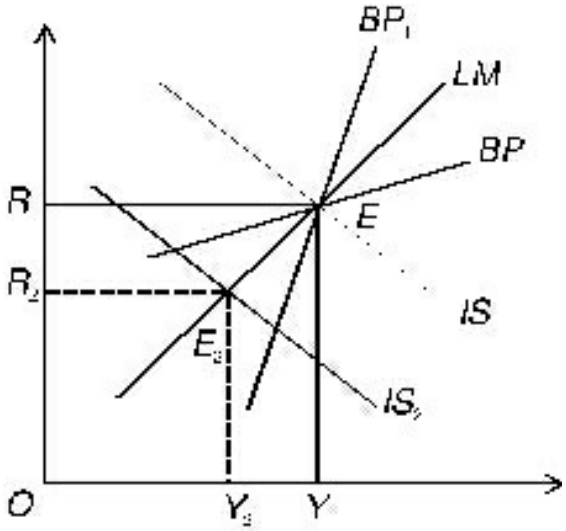


FIG. 3

However, the effects of a contractionary fiscal policy on *BOP* will depend upon the elasticity of the *BP* curve. In the above case, the *BP* curve is *elastic*. If the *BP* curve is *less elastic* such as the BP_1 curve in the figure, a contractionary fiscal policy will lead to a *BOP* surplus. This is because point E_2 is above and to the left of the BP_1 curve.

Expenditure Increasing Fiscal Policy

Take expenditure increasing fiscal policy when the government increases its expenditure or/and reduces taxes. As a result, the *IS* curve shifts upwards to the left as the IS_1 curve which cuts the *LM* curve at E_1, as

shown in Figure 4. This new equilibrium shows increase in interest rate from OR_1 and rise in income from OY to OY_1. The increase in interest rate leads to capital inflow thereby creating short-run *BOP* surplus on capital account. On the other hand, the rise in income increases imports thereby leading to *BOP* deficit on current account. The net overall effect on the *BOP* will depend upon the elasticity of the *BP* curve. If the *BP* curve is *elastic*, as shown, in the figure, and the equilibrium point E_1 is above and to the left of the curve *BP*, there will be overall *BOP* surplus in an *expansionary fiscal policy*. In case the *BP* curve is *less* elastic, shown as the BP_1 curve in the figure, the equilibrium point E_1 being below and to the right of BP_1 curve, there will be overall *BOP* deficit.

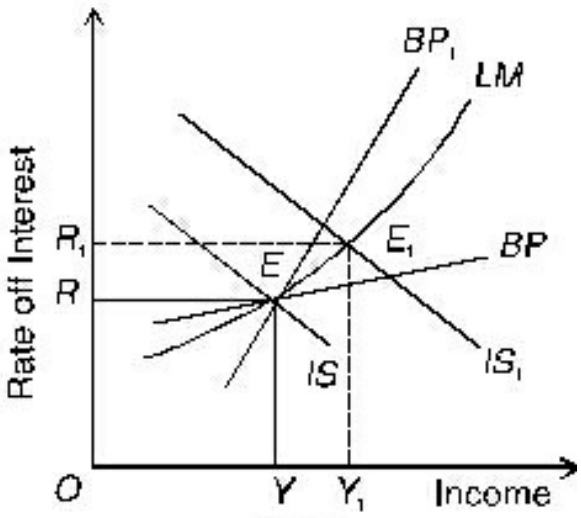


FIG. 4

Conclusion. The above analysis shows that a contractionary monetary policy is effective in correcting a *BOP* deficit. But an expansionary fiscal policy can either improve or worsen a *BOP* deficit. Thus its effects on the overall *BOP* are ambiguous. The increase in interest rate creates short-run *BOP* capital account surplus and the increase in income to a current account deficit. If the former effect predominates an expansionary fiscal policy will bring both the internal and external balance. The same will not be achieved if the latter effect predominates. To solve the *BOP* deficit problem and to bring both the internal and external balance, policy makers suggest the mixing of both monetary and fiscal policies.

MONETARY-FISCAL MIX : INTERNAL AND EXTERNAL BALANCE POLICIES – MUNDELL-FLEMING MODEL

The Mundell-Fleming Model

The Mundell-Fleming model was developed by Fleming* and Mundell**

in the early 1960s for the purpose of analysing the role of monetary and fiscal policies in the context of an open economy. This model is called the *Keynesian Open Economy Model.*

This approach analyses the relationship between two instruments and two targets. The two instruments are monetary policy represented by interest rate and fiscal policy represented by government expenditure. The two objectives or targets are full employment (internal balance) and balance of payments equilibrium (external balance). The assignment rule is to assign monetary policy to the objective of external balance and fiscal policy to internal balance.

Assumptions. The Mundell-Fleming model is based on the following assumptions:

1. Monetary policy is related to change in interest rate.

2. Fiscal policy is related to deficit or surplus budget.

3. Exports are exogeneously given.

4. Imports are a positive function of income.

5. International capital movements respond to domestic interest rate changes.

The Model

Given these Assumptions, Mundell states that "in countries where employment and balance of payments policies are restricted to monetary and fiscal instruments, monetary policy should be reserved for attaining the desired level of the balance of payments, and fiscal policy for preserving internal stability under the conditions assumed here." If monetary and fiscal policies are adjusted smoothly and continuously without long time lags, the assignment rule can work very well. In some cases, it leads straight to the target, while in others it may worsen the other problems temporarily. But ultimately it will achieve the target. This is the *Mundell-Fleming principle of effective use of monetary and fiscal policy*

for internal and external stability, according to which an instrument should be matched with the target on which it exerts the greatest relative influence.

We study below internal and external balance in terms of monetary-fiscal mix policies under fixed and flexible exchange rates with perfect and relative capital mobility and their effects on balance of payments of a country. These are explained as under.

1. Fixed Exchange Rates with Perfect Capital Mobility

When capital is perfectly mobile, a small change in the domestic interest rate brings large flows of capital. The balance of payments is said to be in equilibrium when the domestic interest rate equals the world rate. If the domestic interest rate is lower than the world rate, there will be large capital outflows in order to seek better rates abroad which will be self-eliminating. On the contrary, if the domestic interest rate is higher than the world rate, large capital inflows would bid the domestic interest rate down to its initial level.

---

* R.A. Mundell, *International Economics*, 1968.

** J.M. Fleming, "Domestic Financial Policies under Fixed and Floating Exchange Rates", *I.M.F. Staff Papers*, 1962.

The policy implication of perfect capital mobility is shown in Figure 5 and Figure 6 where the *BP* curve is drawn horizontally. *E* is the initial equilibrium point through which *IS-LM-BP* curves pass. This is the point where *BOP* is zero but the economy is not in full employment equilibrium. This point determines the national income level *OY* and the interest rate *OR*. The *BP* curve is drawn horizontally because even the slightest change in the interest rate will lead to an infinitely large capital flow. If the domestic interest rate is above *OR*, capital flows into the country and if it is below *OR*, capital flows out of the country.

Suppose OY_F is the full employment income level which the economy wants to attain. The monetary authority starts an *expansionary monetary policy* by increasing the money supply. This shifts the *LM* curve to LM_1

which intersects the *IS* curve at E_1 so that the interest rate falls to OR_1. It, in turn, leads to an outflow of capital. Since the price of foreign exchange is fixed, the monetary authority will finance the outflow of capital by selling foreign exchange. The sales of foreign exchange will decrease the money supply. As a result, the LM_1 curve shifts upwards to the left to its original position of the *LM* curve. Thus monetary policy is totally ineffective under fixed exchange rates and perfect international capital mobility in maintaining internal balance. This is because the economy cannot attain the full employment equlibrium point E_2. A contractionary money supply would also be ineffective.

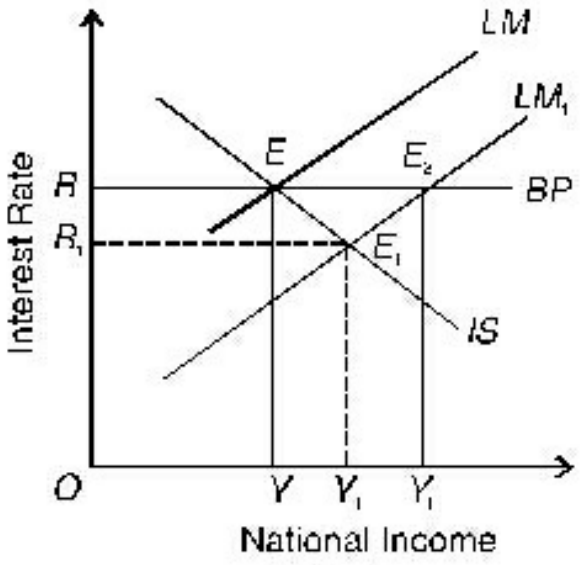


FIG. 5

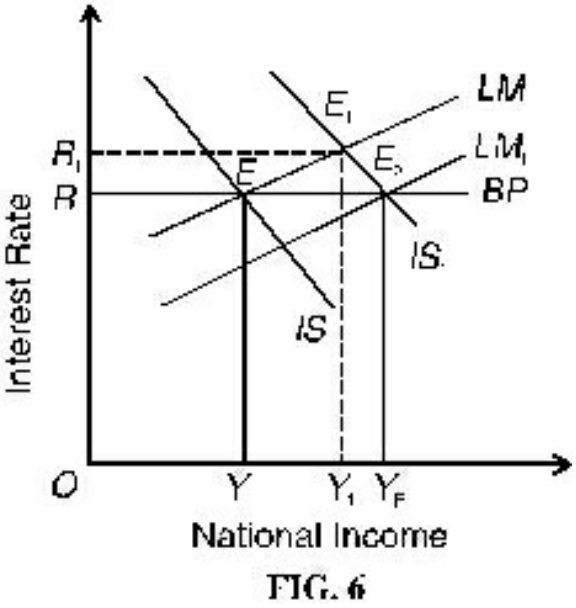


FIG. 6

On the other hand, an *expansionary fiscal policy* has the effect of raising the income level by international capital mobility. This is illustrated in Figure 6. Suppose the government expenditure is increased to achieve full employment level of income OY_F . This shifts the *IS* curve to the right to IS_1 which intersects the *LM* curve at E_1. This causes the interest rate to rise to OR_1 and the income level to fall to OY_1. The rise in interest rate leads to large inflows of capital from abroad. This increases the money supply with the rise in foreign reserves, thereby shifting the *LM* curve to the right to LM_1. Now this curve LM_1 intersects the IS_1 curve at point E_2 where at the fixed exchange rate full employment income level OY_F is reached. Thus fiscal policy by increasing money supply raises aggregate

demand, income and employment.

Thus under perfect capital mobility and fixed exchange rates, *fiscal policy is effective* in maintaining internal balance and *monetary policy is impotent.* So far as the external balance is concerned, it is maintained itself because of perfect capital mobility.

2. Flexible Exchange Rates with Perfect Capital Mobility

These conclusions change when there are flexible exchange rates[4] with perfect capital mobility.

Take an *expansionary monetary policy* which has the effect of lowering the interest rate, increasing capital outflow and thereby bringing deficit in the balance of payments. How this deficit is removed is illustrated in Figure 7. Starting from *E* an expansionary monetary policy shifts the *LM* curve to the right to LM_1 curve, given the *IS* curve. The LM_1 curve intersects the *IS* curve at E_1 which lowers the interest rate to OR_1 and raises income to OY_1. These lead to capital outflows and the consequent deficit in the balance of payments and depreciation of the exchange rate. Depreciation increases the demand for domestic goods in the foreign country, thereby increasing output and income. This moves the economy upward along the LM_1 curve till it reaches point E_2 when income rises to OY_2 and the interest rate rises to the old level *OR*. Equilibrium in the balance of payments is restored at E_2 where the increase in imports through rise in income is offset by surplus in trade balance due to depreciation.

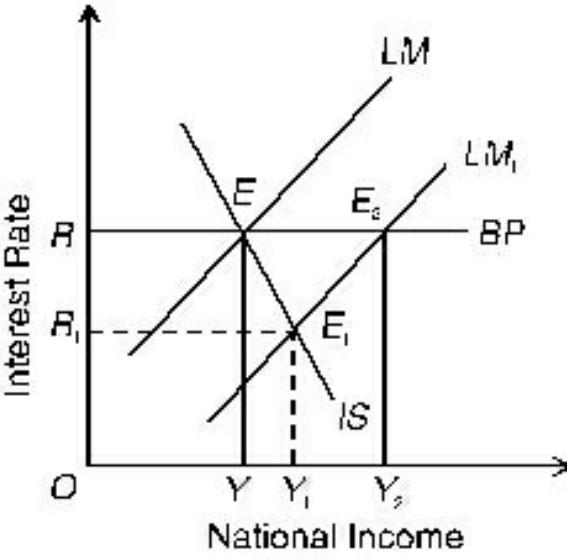


FIG. 7

Take an *expansionary fiscal policy* which shifts the *IS* curve to IS_1, given the *LM* curve in Fig. 8. This brings the economy into equilibrium at E_1

where the IS_1 curve crosses the *LM* curve with OR_1 interest rate and OY_1 income level. Since point E_1 is above the *BP* line, there is surplus in the balance of payments. This surplus leads to the appreciation of the exchange rate which, in turn, reduces the demand for domestic output. This process of appreciation will continue so long as the domestic interest rate is above the world rate and capital inflows continue. Appreciation will continue to reduce the demand for goods and offset the expansionary effect of fiscal policy till the IS_1 curve shifts back to the *IS* curve and the equilibrium is re-established at *E* where the interest rate and the income are back to their original levels of *OR* and *OY*. At *E* the balance of payments is in equilibrium but there is trade deficit due to exchange appreciation which increases the prices of domestic goods for foreigners and reduces the price of imports. Consequently, exports will decline and imports will increase, thereby creating a trade deficit. The equilibrium in the balance of payments is being maintained at E_1 by financing the trade deficit through capital inflows with expansionary fiscal policy. So fiscal policy has no effect on income and employment under perfect capital mobility.

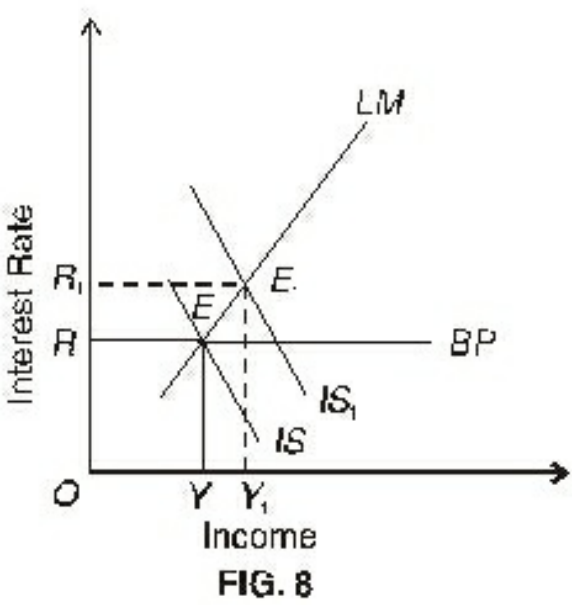


FIG. 8

Thus under flexible exchange rate with perfect capital mobility *monetary policy is effective* in maintaining internal and external balance and fiscal policy is ineffective.

3. Fixed Exchange Rates with Relative Capital Mobility

Consider the effects of monetary and fiscal policies with relative capital mobility under fixed exchange rates. This is illustrated in Figure 9 where the *BP* curve is steeper than the *LM* curve. The initial equilibrium is at point *E* where the curve *IS=LM=BP* curves with *OR* interest rate and *OY* income level. First take an *expansionary fiscal policy* which shifts the *IS*

curve to IS_1 which intersects the *LM* curve at E_2.

---

4. To understand the operation of flexible exchange rates two points are to be noted : (1) When capital outflows, there is depreciation of the currency and when capital inflows, there is appreciation of the currency. (2) Under flexible exchange rate, the demand for foreign exchange equals its supply so that there is neither *BOP* deficit nor surplus. It means that the economy always returns to the original equilibrium position on the *BP* curve when there are changes in income-interest rate combination and shift in the *BP* curve.

The interest rate rises from *OR* to OR_2 and the level of income increases from *OY* to OY_2. There will be a deficit in the balance of payments because E_2 is below and to the right of the *BP* curve. This deficit will bring a decline in the money supply as the monetary authority starts selling foreign exchange. Thus the *LM* curve shifts upward to the left to LM_1, where it intersects the IS_1 and *BP* curves at E_1. Consequently, the economy is at internal and external balance with OR_1 interest rate and OY_1 income level.

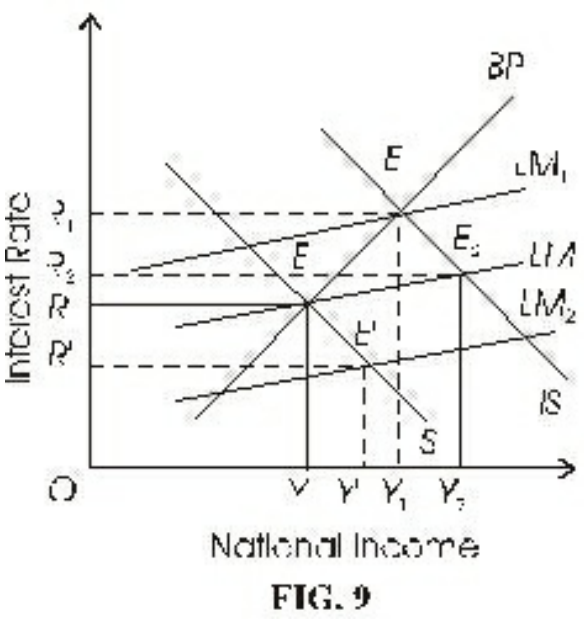


FIG. 9

But under fixed exchange rates with relative capital mobility, an *expansionary monetary policy* will always lead to a deficit in the balance of payments. With this policy, the *LM* curve shifts to the right to LM_2, the interest rate will fall from *OR* to *OR'* and the income will increase from *OY* to *OY'*. The fall in interest rate will lead to a capital account deficit with capital outflow and increase in income to current account deficit with rise in imports. These deficits will force the monetary authority to sell foreign exchange which will reduce the money supply and thus shift the LM_2 curve to its original position of *LM* curve. Consequently, the equilibrium position remains at point *E* with *OR* interest rate and *OY* income level. Thus *monetary policy is ineffective.*

However, *fiscal-monetary policy mix* can lead the economy to both

internal and external balance. This can be achieved by combining an *expansionary fiscal policy with a restrictive monetary policy*. When an expansionary fiscal policy shifts the *IS* curve to the right to IS_1 and a restrictive monetary policy shifts the *LM* curve to the left to LM_1, both intersect the *BP* curve at point E_1 in Fig. 9. Thus the economy attains full employment as well as *BOP* equilibrium with OR_1 interest rate and OY_1 income level which is higher than the original level at point *E*.

4. Flexible Exchange Rates with Relative Capital Mobility

Now consider the effects of monetary and fiscal policy under flexible exchange rates with relative capital movements in terms of Figure 10. First, take the effects of monetary policy with initial equilibrium at point *E* where the curve *BP=IS=LM* curves and *OR* interest rate and *OY* income level. Suppose the monetary authority follows an *expansionary monetary policy* which shifts the *LM* curve to the right to LM_1 and intersects the *IS* curve at E_2 in Fig. 10. This leads to short-run *BOP* deficit because point E_2 is below and to the right of the *BP* curve. With the fall in interest rate to OR_2, there is capital outflow. This leads to increase in the demand for foreign currency and the country's exchange rate depreciates. This increases exports and decreases imports. This causes the *IS* curve to shift to the right to IS_1. *BOP* improves so that the curve shifts to the rights to BP_1. The new equilibrium is established at E_1 where the curve $IS_1=LM_1=BP_1$ curves and both external balance and internal balance are attained at a higher OY_1 income level than *OY*. Thus *monetary policy is effective* under flexible exchange rates. If an *expansionary fiscal policy* is adopted, there is *BOP* deficit under flexible exchange rates. Starting from the equilibrium point *E*, with an increase in

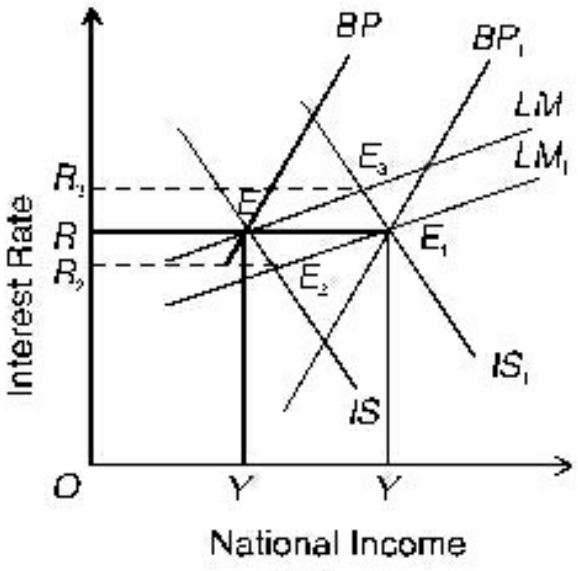


FIG. 10

government expenditure or/and cut in taxes, the *IS* curve will shift to the right to IS_1 which cuts the *LM* curve at E_3. This raises the interest rate to OR_3. There is capital inflow which causes currency appreciation. This, in turn, raises imports and reduces exports and leads to depreciation of the currency, and the IS_1 curve is shifted back to the *IS* curve. The original equilibrium point *E* is reached. Thus *fiscal policy is ineffective* under flexible exchange rates with relative capital mobility. However, an expansionary monetary policy combined with a contractionary fiscal policy under flexible exchange rates and capital movements is effective in attaining internal and external balance.

Monetary and Fiscal Policies for Achieving Internal and External Balance Simultaneously – Swan Model

To achieve the objectives of internal and external balance simultaneously, a judicious combination of expenditure-reducing (internal policies) and expenditure-switching (external policies) instrument is needed. For instance, if the economy is already at the full employment level, a policy of devaluation may cause inflation within the economy. So expenditure-switching policy of devaluation must be accompanied by expenditure-reducing policies of tighter fiscal and monetary controls to maintain full employment and balance of payments equilibrium.

In order to attain simultaneously the two targets of internal and external balance, the relationship between policy instruments are discussed in terms of Trevor Swan's[5] model explained in Figure. 11.

The Swan Model

The Swan model is based on the assumptions that (1) there are no trade restrictions; and (2) there are no capital movements.

Given these assumptions, the model is explained in Fig. 11 where the horizontal axis measures real domestic expenditure, and the vertical axis the exchange rate. A movement to the left (towards *O*) on the horizontal axis means the use of expenditure-reducing policy, and a movement

upwards along the vertical axis means the use of expenditure-switching policy. *IB* is the internal balance curve which represents a situation of full employment and stable prices. It represents the various combinations of exchange rates and real domestic expenditure. The *IB* curve is negatively sloped which shows that an increase in domestic absorption (real expenditure) must be balanced by a decrease in the exchange rate of the country to reduce its trade balance in order to maintain full employment with price stability. Obviously, *points to the right and above the IB curve relate to inflation* or *over full employment*, and *points to the left and below the curve refer to recession or unemployment.*

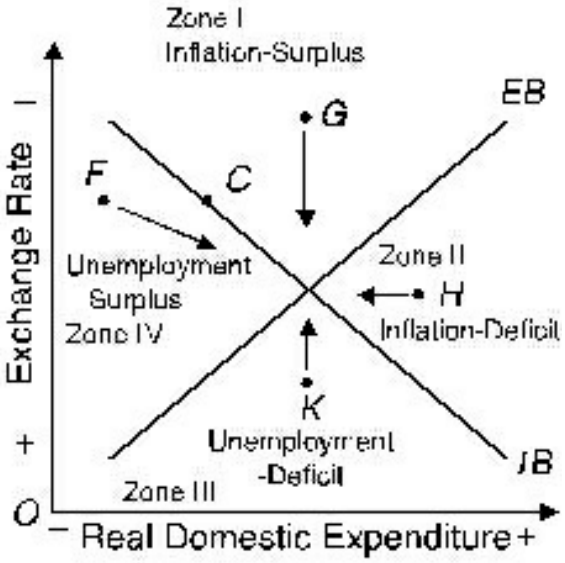


FIG. 11

The *EB* curve represents the external balance where exports equal imports, in the absense of capital movements. So external balance occurs when net exports equal zero. This curve slopes upwards from left to right which shows that for the economy to remain in external balance, devaluation must be balanced by an increase in domestic expenditure (Devaluation will improve the country's trade balance by encouraging exports and discouraging imports, and the increase in real domestic expenditure will increase the country's imports sufficiently). Obviously, *points above the EB curve refer to a surplus and points below the curve relate to a deficit in the balance of payments.*

---

5. *"Long-Run Problems of Balance of Payments,"* in R.E. Caves and H.G. Johnson, *op. cit.*, Ch. 27.

The point where the *IB* curve intersects the *EB* curve represents the point of "bliss" where the economy is simultaneously in internal and external balance. *E* is such a point in Figure 11 where the exchange rate and real domestic expenditure are in equilibrium. If the economy is not at point *E* ,

it is in disequilibrium. According to Swan, “The two curves of internal balance and external balance divide existence into four zones of economic unhappiness.” The four zones of disequilibrium are :

Zone I : Inflation and *BOP* surplus.

Zone II : Inflation and *BOP* deficit.

Zone III : Unemployment and *BOP* deficit.

Zone IV : Unemployment and *BOP* surplus.

To explain the type of policy measures which may be required to achieve internal and external balance simultaneously, we take disequilibrium positions in the four zones.

Take point *G* in Zone I where a *BOP* surplus is combined with inflation. In this situation, the exchange rate should be appreciated to correct the *BOP* surplus and expenditure be reduced to combat inflation. But reduction in expenditure would increase the *BOP* surplus. This represents the “dilemma zone” because no uniform policy can be adopted.

Similarly, point *K* in Zone III where unemployment and *BOP* deficit exist require increase in domestic expenditure to reduce unemployment and depreciation of exchange rate to correct *BOP* deficit. But an expansionary fiscal policy increases income and demand, and thus widens *BOP* deficit. This is again the dilemma zone.

Take point *H* in Zone II where inflation is combined with *BOP* deficit. Inflation should be combated with reduction in domestic expenditure which would also reduce *BOP* deficit along with depreciation of exchange rate. Ultimately, the economy will move towards the equilibrium position *E*.

Finally, move to point *F* in Zone IV where *BOP* surplus is combined with unemployment. Here appreciation of exchange rate and increase in domestic expenditure will move the economy towards internal and external balance at *E*. Zones II and IV are “simple zones”.

Conclusion. The above discussion reveals that if the economy is on neither the *IB* curve nor the *EB* curve, it is in one of the four zones. When the economy follows expenditure changing monetary and fiscal policies simultaneously in dilemma Zones I and III to achieve internal balance (one target), it moves aways from external balance (the other target). To solve this dilemma, Tinbergen developed the rule known as the *Tinbergen Principle* which leads to the assignment problem. The solution to this problem has been suggested by Mundell which we discuss below in detail.

THE ASSIGNMENT PROBLEM : THE MUNDELLIAN MODEL OF MONETARY-FISCAL POLICIES FOR INTERNAL AND EXTERNAL BALANCE

The Assignment Problem

The theory of economic policy has concentrated on two distinct problems. *First*, the relation between the number of policy objectives and the number of policy instruments; and *second*, the assignment of policy instruments to the realisation of targets.

Jan Tinbergen[6] was the first economist to lay down that the number of policy instruments must be equal to the number of objectives. If there are more objectives than policy instruments it means that there are not enough tools to achieve the policy objectives. The system is undetermined. On the other hand, if the number of policy instruments is more than the number of objectives, then there is not one combination of tools and objectives that will solve the problem, but any number. The system is over-determined. Thus the number of policy tools must equal the number of targets for economic policy to be successful. This has come to be known as the *Tinbergen Principle* or the *fixed targets approach.*

In order to achieve given objectives with the same number of policy instruments, the second problem of the assignment of instruments to targets arises. The formulation of the assignments problem will eventually lead to equilibrium values of the objectives, despite lack of co-ordination between them. Thus the assignment probelm relates to the assignment of instruments to targets. The solution to the assignment problem has been suggested by Robert Mundell by the *Principle of Effective Market*

Classification.[7]

The Mundellian Model

Mundell discusses the case of relationship between two instruments and two targets. The two instruments are monetary policy represented by interest rate and fiscal policy represented by government expenditure. The two objectives or targets are full employment (internal balance) and balance of payments equilibrium (external balance). The *assignment rule is to assign monetary policy to the objective of external balance and fiscal policy to internal balance.*

Assumptions. The Mundellian model is based on the following assumptions:

1. Monetary policy is related to changes in interest rate.

2. Fiscal policy is related to deficit or surplus budget.

3. Exports are exogeneously given.

4. Imports are a positive function of income.

5. International capital movements respond to domestic interest rate changes.

The Model

Given these assumptions, Mundell states that "in countries where employment and balance of payments policies are restricted to monetary and fiscal instruments, monetary policy should be reserved for attaining the desired level of the balance of payments, and fiscal policy for preserving internal stability under the conditions assumed here." If monetary and fiscal policies are adjusted smoothly and continuously without long time lags, the assignment rule can work very well. In some cases, it leads straight to the target, while in others it may worsen the other problems temporarily. But ultimately it will achieve the target. This is Mundell's principle of effective use of monetary and fiscal policy for

internal and external stability, according to which an instrument should be matched with the target on which it exerts the greatest relative influence. He calls it the *Principle of Effective Market Classification.*

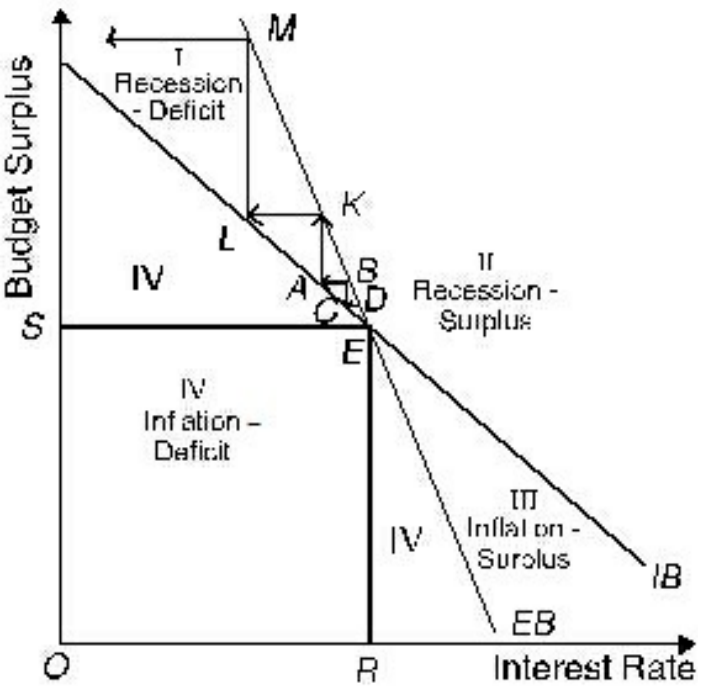


FIG. 12

---

6. J. Tinbergen, *On the Theory of Economic Policy*, 1952.

In Mundell's slightly modified Figure 12, the horizontal axis measures interest rate (monetary policy) and the vertical axis budget surplus (fiscal policy). *IB* is the internal balance line and *EB* the external balance line. The *IB* line represents full employment. It is negatively sloped because a reduction in budget surplus* must be balanced by an increase in interest rate** in order to maintain full employment. There is inflation below this line *IB* (Zone III and IV), and recession above it (Zone I and II). On the other hand, line *EB* gives all points of equilibrium in the balance of payments. It is also negatively sloped because a reduction in the budget surplus must be counteracted by increase in interest rate. There is deficit in the balance of payments below this line (Zone I and IV), and surplus above this line (Zone II and III). The *EB* line is steeper than the *IB* line because an increase in interest rate in order to balance an expansionary fiscal policy (increase in budget deficit or reduction in budget surplus) induces a short-term capital inflow for an external balance. The more responsive capital movements are to interest rate changes, the steeper is the *EB* line relative to the *IB* line. This makes monetary policy relatively more effective for maintaining external balance.

Figure 12 illustrates external and internal balance and the role played by monetary and fiscal policy in maintaining equilibrium between the two at

point *E* with *OR* interest rate and *OS* budget surplus. The two policy measures will take the economy to the equilibrium point *E* in inflation-deficit and recession-surplus situations. But there are policy conflicts in inflation-surplus (Zone III) and recession-deficit (Zone I). Suppose the economy is at point *A* in Zone I where there is full employment within the economy and deficit in the balance of payments. To remove balance of payments deficit, the *monetary authority* acts first by increasing the interest rate by *AB* in order to reduce the money supply. The reduction in money supply will reduce demand for goods and this will, in turn, decrease imports, and restore equilibrium in the balance of payments at *B*. But here the economy is having recession and unemployment. To correct these and to have internal balance, budget surplus will have to be reduced by *BC*. But at *C*, there is again deficit in the balance of payments which necessitates further increase in interest rate by *CD* for reducing the money supply. At *D* the internal balance is again disturbed leading to a further reduction in budget surplus. This process of reduction in money supply followed by reduction in budget surplus will ultimately lead the economy to the equilibrium point *E* where there is simultaneous internal and external balance. Thus the use of monetary policy for external balance and fiscal policy for internal balance will lead to equilibrium in Zones II and IV.

---

7. R.A. Mundell, "The Appropriate Use of Fiscal and Monetary Policy for Internal and External Stability," *I.M.F. Staff Papers,* 1962. *International Economics*, 1968.

* Reduction in budget surplus means increase in government expenditure. It is contractionary fiscal policy.

On the other hand, if *budget surplus is used to remedy the deficit in balance of payments and monetary policy to correct recession and unemployment, there would be neither external balance nor internal balance*. Starting from point *A*, an increase in budget surplus would move the economy to *K* where the external balance is achieved but there is recession and unemployment in the economy. To remedy it, the interest rate is reduced by *KL* for increasing the money supply. But at *L* deficit in the balance of payments rises over its previous level. This will require a still greater budget surplus by *LM*. This will necessitate still larger reduction in interest rate to remove recession and unemployment. In this

way, the economy would move further and further away from point *E* and there would not be simultaneous internal and external balance. In this case, the assignment rule leads to explosive instability because the two policies are badly coordinated.

In fact, Mundell argues for a judicious monetary and fiscal-policy mix. Both objectives will be realised, when monetary policy is paired with the objective of external balance and fiscal policy with the objective of internal balance.

Criticisms of Mundell's Model

There are several shortcomings of Mundell's analysis.

1. Unrealistic Assumptions. This model assumes that the authorities know about the *IB* and *EB* curves, the zone in which the economy is operating, and the extent to which the economy is away from both internal and external balance so that appropriate monetary and fiscal policy can be applied. It also presupposes that they know the quantitative results which are expected from the application of each policy. But it is not possible to accurately estimate the degree of disequilibrium due to lack of data about them. Accordingly, the policy changes may not be appropriate to the type of disequilibrium.

2. Ignores Stagflation. This analysis overlooks the situation of unemployment and inflation. This is unrealistic because this phenomenon, known as stagflation, is found in almost all developed countries.

3. Neglects Other Factors. This analysis considers only differences in interest rates as the cause of capital movements and neglects other factors such as exchange rate variations. Moreover, it is not possible that a persistent deficit may be financed by means of capital movements.

4. Practical Constraints. Monetary and fiscal policies operate under certain practical constraints. Due to political reasons, some governments are unable to follow a restrictive fiscal policy and a monetary policy of high interest rates. Even if such policies can be started, they may not be successful because capital flows may not be interest-sensitive.

5. Cannot Correct Current Account Deficit. The prescribed policy mix may be unable to correct a current account deficit. Since the policy mix affects both the capital flows and imports, it can only ensure that a negative trade balance is offset by a positive capital flow, and vice versa.

6. Not a True Adjustment Mechanism. The monetary-fiscal mix is not a true adjustment mechanism. It is just a palliative. It does not adjust the balance of payments but simply stabilises it. Capital flows only fill the gap between autonomous demand and supply of foreign exchange, leaving prices and incomes unchanged.

** Increase in interest rate implies reduction in the money supply. It is contractionary monetary policy.

7. Debt-Servicing Requirements not Considered. This analysis does not take into account the debt-servicing requirements that a continuous capital inflow would have on the current account of the balance of payments when the domestic interest rate is raised.

8. Retards Capital Formation. When the interest rate is raised through monetary policy, it will lead to a decrease in investment at home. This must be accompanied either by an increase in government expenditure or by tax reductions or by a combination of both. Such a monetary-fiscal mix wastes the economy's savings by diverting them into debt-financed government expenditure which retards capital formation. According to Johnson, this raises the problem of "inefficiency versus efficiency in the use of domestic saving potential."

9. Conflicting Policies. There is the possibility of conflicts between the prescribed policy mixes among governments of different countries. According to Johnson, it is a difficult and highly complicated process to arrive at the right combination of monetary and fiscal policies in all countries simultaneously. If, however, such a combination is worked out by trial and error, it may lead away from rather than toward equilibrium.

10. Long Time Lags. The model assumes that there are no long time lags for the operation of monetary-fiscal policies. But both monetary and fiscal lags are quite long and they retard the process of simultaneous equilibrium

of internal and external balance.

EXPENDITURE SWITCHING POLICIES

Expenditure switching policies refer to devaluation or revaluation of a country's currency in order to switch its expenditure from foreign to domestic goods or vice versa. They aim at correcting *BOP* disequilibrium. But Johnson distinguishes between two types of expenditure switching policies. The *first* is devaluation, and the *second* is the use of *direct controls* to restrict imports and to correct *BOP* deficit. We shall follow Johnson in explaining expenditure switching policies. Thus expenditure switching policies aim at increasing the demand for domestic goods and to switch expenditure from imported to domestic goods. Expenditure switching policies aim at maintaining external balance.

1. Devaluation

Devaluation is referred to as expenditure switching policy because it switches expenditure from imported to domestic goods and services. Devaluation means a reduction in the external value of a currency in terms of other currencies. But there is no change in the internal purchasing power of the country. Thus when a country with *BOP* deficit devalues its currency, the domestic price of its imports increases and the foreign price of its exports falls. This makes its exports cheaper and imports dearer. Now the foreigners can buy more goods by paying less money than before devaluation. This encourages exports. This causes expenditures to be switched from foreign to domestic goods as the country's exports increase and the country produces more to meet the domestic and foreign demand for goods. On the other hand, with imports becoming dearer than before, they decline. Thus with the rise in exports and fall in imports, *BOP* deficit is corrected.

Assumptions. This analysis is based on the following assumptions :

1. The elasticity of demand for exports and imports is elastic.

2. The supply of exports is sufficient to meet the increased demand for

exports after devaluation.

3. The internal price level remains constant after devaluation.

4. The other country does not devalue its currency simultaneously.

5. The other country does not adopt such counter-devaluation measures as levying tariff duties on the exports of the devaluing country.

Explanation

Give these assumptions, the effects of devaluation on *BOP* deficit of a country are explained in terms of Figure 13.

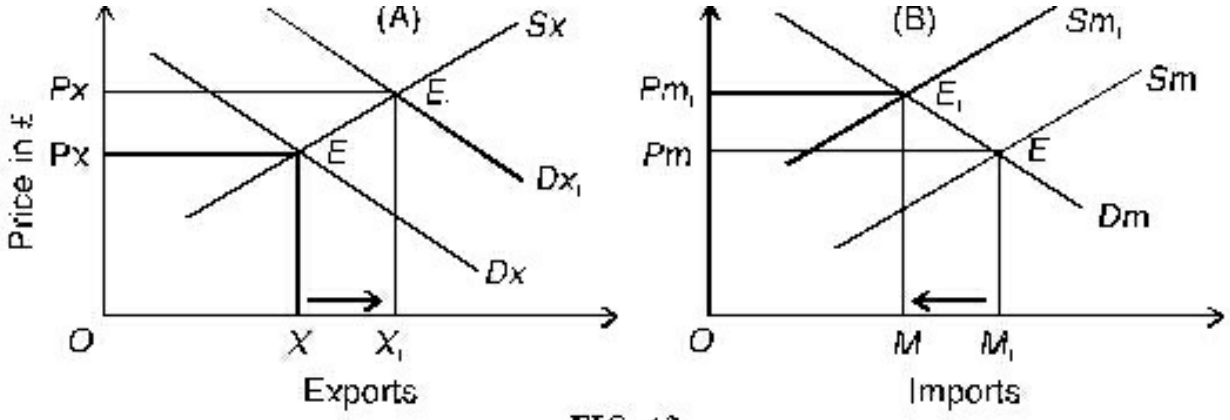


FIG. 13

In order to assess the true effect of devaluation on *BOP* deficit, it is advisable to study price movements in exports and imports in the same currency. Figure 13 (A) and (B) illustrates these effects of devaluation on exports and imports respectively where *Dx* and *Sx* are the demand and supply curves of exports and *Dm* and *Sm* are the demand and supply curves of imports. Suppose the British pound (£) is devalued in relation to the U.S. dollar ($) and the price movements before and after devaluation are taken in pound. Both the demand and supply curves of exports and imports are taken as elastic. *First,* take exports and Panel (A) of the figure. Before devaluation, Britain exports *OX* quantity at OP_x price to the U.S. Devaluation of pound has no effect on the supply of exports in pound. Therefore, the supply curve of exports S_X does not change. But to the U.S. consumers of British goods, devaluation of pound means cheaper goods

than before. Consequently, the demand for exports increases and the demand curve for export *Dx* shifts to the right to Dx_1. The pre-devaluation price of *OX* exports is OP_x. After devaluation of the pound, the export price rises to OPx_1 and the volume of exports increases to OX_1.

Now take the effect of devaluation on imports. Before devaluation, Britain imports OM_1 quantity at *OPm* price from the U.S. With devaluation, imports become dearer in pound and their volume is reduced than before devaluation. Therefore, the supply curve of imports *Sm* shifts to Sm_1 in Panel (B) of the figure. But the demand curve for imports *Dm* being elastic, the increase in the price of imports from *OPm* to OPm_1 reduces the quantity bought from OM_1 to *OM*.

Thus by increasing exports by XX_1 and decreasing imports by M_1M, devaluation in terms of the currency of the devaluing country brings equilibrium in the balance of payments.

Limitations. The following are the limitations of devaluation as a measure to correct *BOP* deficit of a country:

1. Less Elastic Demand. The above analysis assumes an elastic demand for exports and imports. In reality, the demand for export and imports is never more than unity. Rather, it is always less than unity. As a result, the effect of devaluation on *BOP* will be unfavourable.

2. Supply Shortages. It also assumes an adequate supply of exports to meet the increased demand for exports after devaluation. But it is not possible for a devaluing country to increase the supply of exports due to the scarcity of domestic raw materials, lack of incentives to manufacturers and exporters of exportable goods, reduction in export duties, etc.

3. Price Rise. It is further assumed that the internal price level in the devaluing country remains constant. This is unrealistic because when the demand for exports increases, there is a shortage of such goods in the domestic market which raises their prices. Similarly, with devaluation, prices of imported goods increase. Thus, the rise in the prices of goods

nullify the favourable effects of devaluation.

4. Other Countries Devalue. It is presumed that other country does not devalue its currency. But there is every likelihood that the other country, whose exports are adversely affected may retaliate and devalue its currency. Consequently, the favourable effects of devaluation will be neutralised.

5. Counter-Devaluation Measures. If other countries adopt trade restriction measures like raising of tariff duties and other direct controls, devaluation will fail to reduce *BOP* deficit.

Despite these limitations, devaluation is regarded as a potent device to control *BOP* deficit by countries.

2. Direct Control

The second type of expenditure switching policy is the use of direct controls to restrict imports of goods in order to correct a *BOP* deficit. Such a policy increases domestic output for export and encourages production of import substitute goods. This policy encourages consumers to purchase domestic substitutes and domestic producers to produce import substitutes. To induce producers to switch their expenditures to exportable goods, the government may give them production and export subsidies. It may also restrict import of undesirable or unimportant items by levying heavy import duties, fixation of quotas, etc. At the same time, it may allow imports of essential goods duty free or at lower import duties, or fix liberal import quotas for them. For instance, the government may allow free entry of capital goods, but impose heavy import duties on luxuries. Import quotas are also fixed and the importers are required to take licenses from the authorities in order to import certain essential commodities in fixed quantities. In these ways, imports are reduced in order to correct an adverse balance of payments. Johnson calls them *commercial controls* which operate on the goods side of transactions by preventing people from buying certain goods or forcing them to buy others, or providing financial incentives like tariff subsidies, etc. for certain kinds of sales or purchases.

The government also adopts *financial controls* to reduce a *BOP* deficit. They "operate through control over the use of money, by restricting the freedom of the use of domestic money either through regulation of certain uses (as in the case of multiple exchange rates) or by making some uses of money more expensive than others." These are stringent exchange control measures. Exchange controls have a dual purpose. They restrict imports and regulate foreign exchange. They may include full control over all foreign exchange receipts and payments by the monetary authority. Foreign currencies are required to be surrendered to exchange control authorities. There may be restrictions on sale and purchase of foreign currencies and securities; on direct investments abroad on short-term speculative capital outflows; on royalties, interest and amortisation payments to foreigners on foreign travels, on tourist expenditure abroad, etc. The government may also resort to multiple exchange rates. It may use a lower exchange rate for essential imports such as raw materials, machinery, etc., and a higher exchange rate for non-essential items.

Thus direct controls, both commercial and financial, by reducing imports and regulating foreign exchange for the needs of the economy help in correcting an adverse *BOP*.

Limitations. But direct controls involve large social costs. They lead to welfare losses when people are prevented from using foreign exchange and import goods.They also involve large administrative costs. There is also wastage of resources when people try to evade the exchange controls and apply for foreign exchange licenses which are auctioned. There may also be retaliation from other countries which may offset its benefits, if any. Morever, direct controls deal with *BOP* disequilibrium by suppressing the deficit that is a symptom, and not the basic trouble.When the deficit is due to capital flight, suppressing the deficit does not cure the underlying cause. Direct controls deal with only the deficit and fail to come to grips with the basic cause. Thus direct controls offer temporary rather than a permanent solution to a *BOP* deficit. That is why the majority of countries favour a liberal trade regime where monetary-fiscal measures are considered better than direct controls in correcting a *BOP* deficit. Despite its limitations, even devaluation is considered better than direct controls.

EXERCISES

1. Indicate the way and the extent to which monetary and fiscal instruments could be successfully used to bring about internal and external balance.
2. What policy instruments would you suggest for maintaining both internal and external equilibrium?
3. Discuss Mundell-Fleming Medel under fixed and flexible exchange rates.
4. Explain the Swan Model for achieving internal and external balance.
5. Distinguish between expenditure switching and expenditure changing policies for balance of payments adjustment. Discuss the effectiveness of devaluation in correcting a balance of payments deficit.
6. Discuss the various types of direct controls with which a balance of payments disequilibrium can be corrected.
7. Critically examine the role of expenditure reducing policies in correcting a deficit in balance of payments.
8. What do you mean by the Assignment Problem? How has it been solved by the Mundellian model? Explain critically.

CHAPTER 62

FOREIGN EXCHANGE RATE

MEANING OF FOREIGN EXCHANGE RATE

The foreign exchange rate or exchange rate is the rate at which one currency is exchanged for another. It is the price of one currency in terms of another currency. It is customary to define the exchange rate as the price of one unit of the foreign currency in terms of the domestic currency. The exchange rate between the dollar and the pound refers to the number of dollars required to purchase a pound. Thus the exchange rate between the dollar and the pound from the US viewpoint is expressed as $ 2.50 = £ 1. The Britishers would express it as the number of pounds required to get one dollar, and the above exchange rate would be shown as £ 0.40 = $ 1.

The exchange rate of $ 2.50 = £ 1 or £ 0.40 = $ 1 will be maintained in the world foreign exchange market by arbitrage. *Arbitrage refers to the purchase of a foreign currency in a market where its price is low and to sell it in some other market where its price is high*. The effect of arbitrage is to remove differences in the foreign exchange rate of currencies so that there is a single exchange rate in the world foreign exchange market. If the exchange rate is $ 2.48 in the London exchange market and $ 2.50 in the New York exchange market, foreign exchange speculators, known as arbitrageurs, will buy pounds in London and sell them in New York, thereby making a profit of 2 cents on each pound. As a result, the price of pounds in terms of dollars rises in the London market and falls in the New York market. Ultimately, it will equal in both the markets and arbitrage comes to an end. If the exchange rate betwen the dollar and the pound

rises to $ 2.60 = £ 1 through time, the dollar is said to depreciate with respect to the pound, because now more dollars are needed to buy one pound. When the rate of exchange between the dollar and the pound falls to $ 2.40 = £ 1, the value of the dollar is said to appreciate because now less dollars are required to purchase one pound. If the value of the first currency depreciates that of the other appreciates, and vice versa. Thus a depreciation of the dollar against the pound is the same thing as the appreciation of the pound against the dollar, and vice versa.

DETERMINATION OF EQUILIBRIUM EXCHANGE RATE

The exchange rate in a free market is determined by the demand for and the supply of foreign exchange. The equilibrium exchange rate is the rate at which the demand for foreign exchange equals to supply of foreign exchange. In other words, it is the rate which clears the market for foreign exchange. Ragner Nurkse defined the equilibrium exchange rate as, “that rate which over a certain period of time, keeps the balance of payments in equilibrium.” There are two ways of determining the equilibrium exchange rate. The rate of exchange between dollars and pounds can be determined either by the demand and supply of dollars with the price of dollars in pounds, or by the demand and supply of pounds with the price of pounds in dollars. Whatever method is adopted, it yields the same result. The analysis that follows is based on the dollar price in terms of pounds.

The Demand for Foreign Exchange

The demand for foreign exchange is a derived demand from pounds. It arises from import of British goods and services into the US and from capital movements from the US to Britain. In fact, the demand for pounds implies a supply of dollars. When the US businessmen buy British goods and services and make capital transfers to Britain, they create demand for British pounds in exchange for US dollars because they cannot make payments to Britain in their currency, the US dollars.

The demand curve for pounds *DD* is downward sloping from left to right in Figure 1. It implies that the lower the exchange rate on pounds, the

larger will be the quantity of pounds demanded in the foreign exchange (US) market, and vice versa. This is because a lower exchange rate on pounds make British exports of goods and services cheaper in terms of dollars. The opposite happens if the exchange rate on pound is higher. It will make British goods and services dearer in terms of dollars, and the demand for pounds will fall in the foreign exchange (US) market.

But the shape of the demand curve for foreign exchange will depend on the elasticity of demand for imports. "If a country imports necessities and raw materials, we may expect the elasticity of demand for imports to be low and the quantity imported to be insensitive to price changes. If, on the other hand, the country imported luxury goods and goods for which suitable substitutes exist, demand elasticities for imports might be high... . If the country has many well-developed import competing industries, the elasticity of demand for imports most certainly is high... . In the short run, elasticity of demand for imports may not be very high. In the long run, however it is much more probable that the production pattern will alter according to price changes, and the demand for imports, therefore, will be more elastic."[1]

The Supply of Foreign Exchange

The supply of foreign exchange in our case is the supply of pounds. It arises from the US exports of goods and services and from capital movements from the US to Britain. Pounds are offered in exchange for dollars because British holders of pounds wish to make payments in dollars. Thus the supply of foreign exchange reflects the quantities of pounds that would be supplied in the foreign exchange market at various dollars prices of pounds.

The supply curve for pounds *SS* is an upward sloping curve, as shown in Figure 1. It is a positive function of the exchange rate on pounds. As the exchange rate on pounds increases, the greater is the quantity of pounds supplied in the foreign exchange market. This is because with increase in the dollar price of pounds (lower pounds price of dollars), US goods, services and capital funds become better bargains to holders of pounds. Therefore the holders of pounds will offer larger quantities of pounds with the increase in the exchange rate.

But the shape of supply curve of foreign exchange will be determined by the elasticity of the supply curve. "As the value of the country's own currency increases, imports become relatively cheaper, and more is imported. As more is imported, more of the home currency is supplied on the foreign exchange market, provided elasticity is greater than unity. When imports become relatively cheap, new goods will start to be imported, and domestic import-competing industry will be gradually eliminated by imports. These are two important reasons why we expect the supply of foreign exchange to be quite elastic. Further, the larger the time perspective we take into account, the more elastic will be the supply."[2]

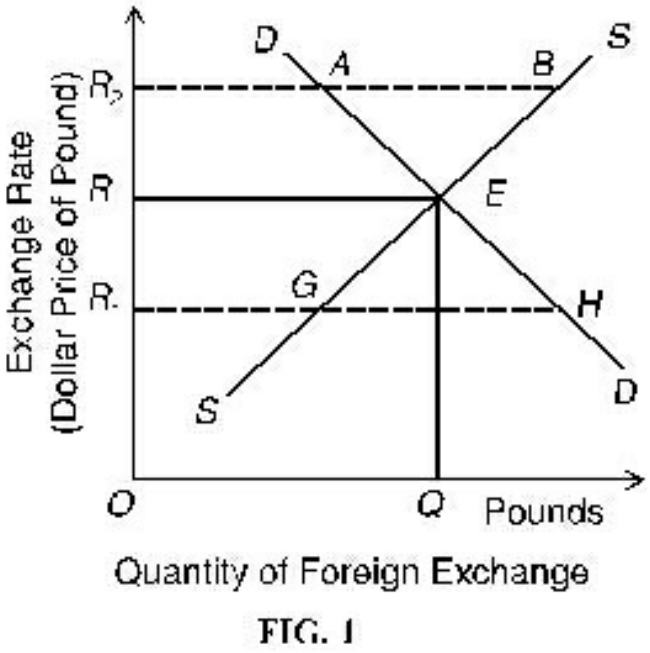


FIG. 1

Equilibrium Exchange Rate

Given the demand and supply curves of foreign exchange, the equilibrium exchange rate is determined where *DD*, the demand curve for pounds intersects *SS*, the supply curve of pounds. They cut each other at point *E* in Figure 1. The equilibrium rate is *OR* and *OQ* of foreign exchange is demanded, and supplied. At *OR* exchange rate the US demand for pounds equals the British supply of pounds, and the foreign exchange market is cleared. At any higher rate than this, the supply of pounds would be larger than the demand for pounds so that some people who wish to convert pounds into dollars will be unable to do so. The price of pounds will fall, less pounds will be supplied and more will be demanded. Ultimately, the equilibrium rate of exchange will be re-established. In Figure 1 when the exchange rate increases to OR_2, the supply of pounds $R_2B > R_2A$ the demand for pounds. With the fall in the price of pounds, the equilibrium exchange rate OR_2 is again established at point *E*. On the contrary, at an

exchange rate lower than this, say OR_1, the demand for pounds R_1H is greater than the supply of pounds R_1G. Some people who want pounds will not be able to get them. The price of pounds will rise which will reduce the demand and increase the supply of pounds so that the equilibrium exchange rate OR is re-established at point E where the two curves DD and SS intersect.

---

1. Bo Sodersten, *op. cit.,* p. 211. Students who find it difficult to understand may leave this para without loss in continuity.

2. *Ibid.,* 214-15. Students who find it difficult may also leave this para.

Suppose there is a shift upward in the US demand for pounds, as shown by the upward shifting of the DD curve to D_1D_1 in Figure 2 (A). This may be due to increase in the US tastes for British goods, an increase in the US national income, etc. which increases the demand for imported goods in the US. With the shifting up of the demand curve to D_1D_1, the US dollar depreciates and the British pound appreciates which re-establish the new equilibrium exchange rate OR_2 at point E_2 where OQ_2 quantity of foreign exchange is demanded and supplied.

On the other hand, if the supply of pounds increases and the supply curve shifts down from SS to S_1S_1, as shown in Figure 2 (B), the value of pounds depreciates and that of dollars appreciates. This automatically brings about a new equilibrium exchange rate OR_1 at point E_1 in Figure 2 (B) where the S_1S_1 curve intersects the DD curve. At the new equilibrium exchange rate OR_1, OQ_1 of foreign exchange is demanded and supplied. The supply of pounds may increase due to the increase in the tastes of Britishers for the US goods, the increase in the national income of Britain, etc.

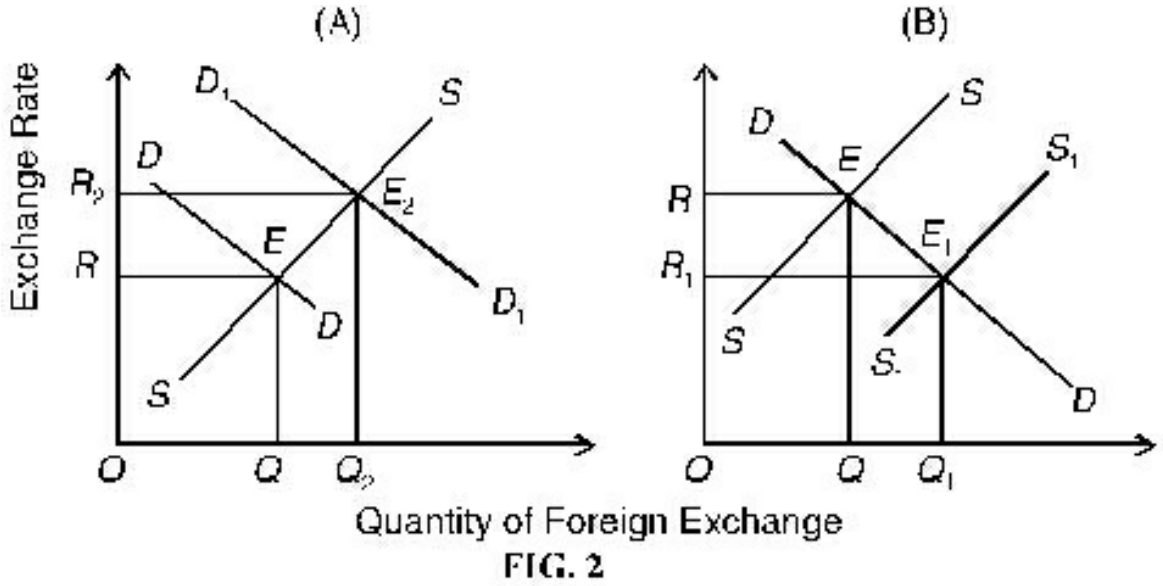


FIG. 2

Thus under flexible exchange rates, equilibrium rate of exchange will prevail which will clear the market and keep the balance of payments in equilibrium.

THEORIES OF FOREIGN EXCHANGE RATE

There are three theories of the determination of foreign exchange rate. The first is the Mint Parity Theory, the second is the Purchasing Power Parity Theory, and the third is the Balance of Payments Theory. We discuss these theories one by one.

1. The Mint Parity Theory : Determination Under Gold Standard

This theory is associated with the working of the international gold standard. Under this system, the currency in use was made of gold or was convertible into gold at a fixed rate. The value of the currency unit was defined in terms of certain weight of gold, that is, so many grains of gold to the rupee, the dollar, the pound, etc. The central bank of the country was always ready to buy and sell gold at the specified price. The rate at which the standard money of the country was convertible into gold was called the *mint price* of gold. If the official British price of gold was £6 per ounce and the US price of gold $ 36 per ounce, they were the mint prices of gold in the respective countries. The exchange rate between the

dollar and the pound would be fixed at $ 36/ £6 = $ 6. This rate was called the mint parity or mint par of exchange because it was based on the mint price of gold. Thus under the gold standard, the normal or basic rate of exchange was equal to the ratio of their mint par values (R = $/£).

But the actual rate of exchange could vary above and below the mint parity by the cost of shipping gold between the two countries. To illustrate this, suppose the US has a deficit in its balance of payments with Britain. The difference between the value of imports and exports will have to be paid in gold by US importers because the demand for pounds exceeds the supply of pounds. But the transhipment of gold involves transportation cost and other handling charges, insurance, etc. Suppose the shipping cost of gold from the US to Britain is 3 cents. So the US importers will have to spend $ 6.03 ($ 6 + .03c) for getting £ 1. This could be the exchange rate which is the US *gold export point* or *upper specie point.* No US importer would pay more than $ 6.03 to obtain one pound because he can buy $ 6 worth of gold from the US treasury and ship it to Britain at a cost of 3 cents per ounce. Similarly, the exchange rate of the pound cannot fall below $ 5.97 in the case of a surplus in the US balance of payments. Thus the exchange rate of $ 5.97 to a pound is the US *gold import point* or *lower specie point.*

Assumptions. This theory is based on the following assumptions:

1. The price of gold is fixed by a country in terms of its currency.

2. It buys and sells gold in any amount at that price.

3. Its supply of money consists of gold or paper currency which is backed by gold.

4. Its price level varies directly with its money supply.

5. There is movement of gold between countries.

6. Capital is mobile within countries.

7. The adjustment mechanism is automatic.

Explanation

Given these assumptions, the exchange rate under the gold standard is determined by the forces of demand and supply between the gold points and is prevented from moving outside the gold points by shipments of gold.

Figure 3 shows the determination of the exchange rate under the gold standard. The exchange rate *OR* is set up at point *E* where the demand and supply curves DD_1 and SS_1 intersect. The exchange rate need not be at the mint parity. It can be anywhere between the gold points depending on the shape of the demand and supply curves. The mint parity is simply meant to define the US gold export point ($ 6.03) and the US gold import point ($ 5.97). Since the US treasury is prepared to sell any quantity of gold at a price of $ 36 per ounce, no American would pay more than $ 6.03 per pound, because he can get any quantity of pounds at that price by exporting gold. That is why, the US supply curve of pounds becomes perfectly elastic or horizontal at the US gold export point. This is shown by the horizontal portion S_1 of the SS_1 supply curve. Similarly, as the US treasury is prepared to buy any quantity of gold at $ 36 per ounce, no American would sell pounds less than $ 5.97 because he can sell any quantity of pounds at the price by gold imports. Thus the US demand curve for pounds becomes perfectly elastic at the US gold import point. This is shown by the horizontal portion D_1 of the demand curve DD_1.

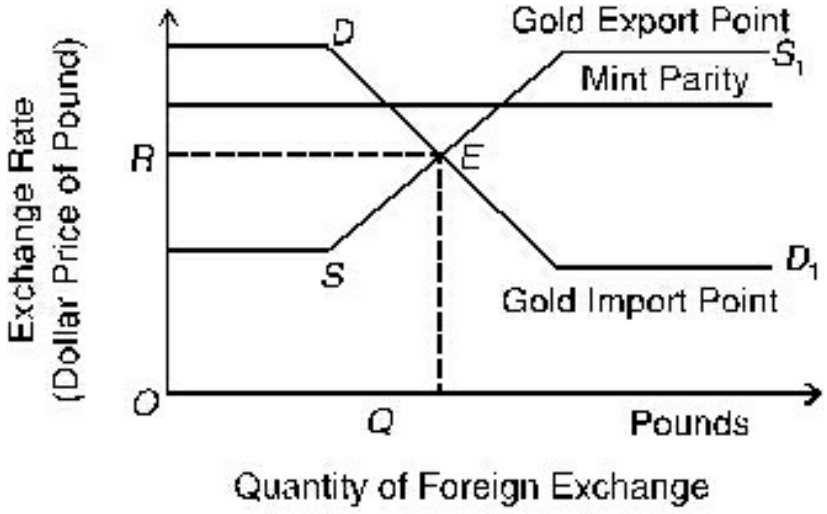


FIG. 3

Criticisms. The mint parity theory has been criticised on the following grounds :

1. The international gold standard does not exist now ever since it broke down after the Depression of the 1930s.

2. The theory is based on the assumption of free buying and selling of gold and its movement between countries. But governments do not allow such sales, purchases and movements.

3. The theory fails to explain the determination of exchange rates as most countries are on inconvertible paper standard.

4. This theory assumes flexibility of internal prices. But modern governments follow independent domestic price policy unrelated to fluctuations in exchange rate.

Conclusion. The mint parity theory has long been discarded ever since the gold standard broke down. No country is on the gold standard now. There are neither free movements of gold nor gold parities. So this theory has only an academic interest.

2. The Purchasing Power Parity Theory

The purchasing power parity (PPP) theory was developed by Gustav Cassel in 1920 to determine the exchange rate between countries on inconvertible paper currencies. The theory states that equilibrium exchange rate between two inconvertible paper currencies is determined by the equality of the relative change in relative prices in the two countries. In other words, the rate of exchange between two countries is determined by their relative price levels.

There are two versions of the PPP theory : the absolute and the relative. The absolute version states that the exchange rate between two currencies should be equal to the ratio of the price indexes in the two countries. The formula is $R_{AB} = P_A/P_B$ where R_{AB} is the exchange rate between two countries A and B and P refers to the price index. This version is not used because it ignores transportation costs and other factors which hinder trade, non-traded goods, capital flows and real purchasing power. Economists, therefore, use the relative version which we discuss.

The theory can be explained with the help of an example.

Suppose India and England are on inconvertible paper standard and by spending Rs. 60, the same bundle of goods can be purchased in India as can be bought by spending £ 1 in England. Thus according to the purchasing power parity theory, the rate of exchange will be Rs. 60 = £ 1.

If the price levels in the two countries remain the same but the exchange rate moves to Rs. 50 = £ 1. This means that less rupees are required to buy the same bundle of goods in India as compared to £ 1 in England. It is a case of *overvaluation* of the exchange rate. This will encourage imports and discourage exports by India. As a result, the demand for pounds will increase and that of rupees will fall. This process will ultimately restore the normal exchange rate of Rs. 60 = £ 1. In the converse case, if the exchange rate moves to Rs. 70 = £ 1, the Indian currency becomes *undervalued.* As a result, exports are encouraged and imports are discouraged. The demand for rupees will rise and that for pounds will fall so that the normal exchange rate of Rs. 60 = £ 1 will be restored.

According to the theory, the exchange rate between two countries is determined at a point which expresses the equality between the respective purchasing powers of the two currencies. This is the purchasing power parity which is a moving par and not fixed par (as under the gold standard). Thus with every change in price level, the exchange rate also changes. To calculate the equilibrium exchange rate, the following formula is used :

$$R = \frac{\text{Domestic Price of a Foreign Currency} \times \text{Domestic Price Index}}{\text{Foreign Price Index}}$$

$$\text{OR } R = R_0 \times \frac{P_{A_1} / P_{A_0}}{P_{B_1} / P_{B_0}}$$

where 0 = base period, 1 = period 1, *A* and *B* countries, *P* = price index and R_0 = exchange rate in base period.

According to Cassel, the purchasing power parity is "determined by the quotients of the purchasing powers of the different currencies." This is what the formula does. Let us explain it in terms of our above example. Before the change in the price level, the exchange rate was Rs. 60 = £ 1. Suppose the domestic (Indian) price index rises to 300 and the foreign (England) price index rises to 200, thus the new equilibrium exchange rate will be

$$R = £\ 1 \times \frac{300}{200} = £\ 1.5$$

or Rs. 60 = £ 1.5

This will be the purchasing power parity between the two countries. In reality, the parity will be modified by the cost of transporting goods including duties, insurance, banking and other charges. These costs of transporting goods from one country to another are, in fact, the limits within the exchange rate can fluctuate depending upon the demand and supply of a country's currency. There is the upper limit, called the commodity export point; and the lower limit, known as the commodity import point. (These limits are not as definite as the gold points under the mint par theory).

The PPP theory is illustrated in Figure 4 where *DD* is the demand curve for foreign currency (pound in our example) and *SS* is the supply

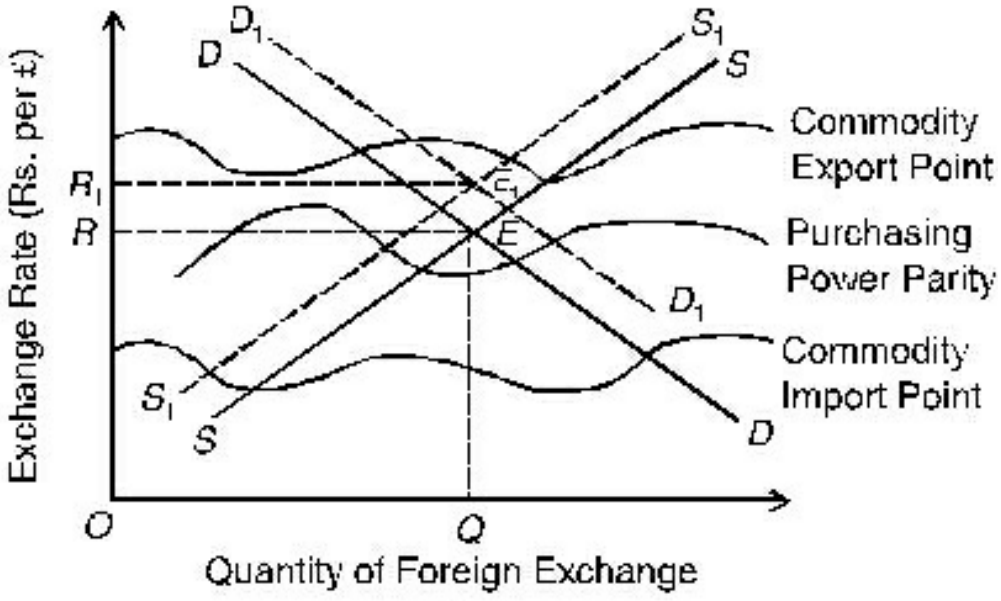


FIG. 4

curve of currency. *OR* is the rate to exchange of rupees per £, which is determined by their intersection at point *E* so that the demand for the supply of foreign exchange equals *OQ* quantity. Suppose the price level rises in India and remains constant in England. This makes Indian exports costly in England and imports from England relatively cheaper in India. As a result, the demand for pounds increases and the supply of pounds decreases. Now the *DD* curve shifts upward to the right to D_1D_1 and the *SS* curve to the left to S_1S_1. The new equilibrium exchange rate is set at OR_1 rupees per pound, which represents the new purchasing power parity. The exchange rate rises by the same percentage as the Indian price level. The purchasing power curve shows that with relative change in the price levels, the exchange rate tends to fluctuate along this curve above or below the normal exchange rate. But there is a limit upto which the purchasing power parity curve can move up and down. The upper and lower limits are set by the commodity export point and the commodity import point respectively.

Criticisms. Cassell's PPP theory became very popular among economists during 1914-24 and was widely accepted as a realistic explanation of the determination of foreign exchange rate under inconvertible paper currencies. But it has been severely criticised for its weak theoretical base. Some of the criticisms are discussed as under :

1. Defects in Calculating Price Level. One of the serious defects of the theory is that of calculating the price levels in the two countries. The use of index number in calculations presents many difficulties such as the base year, coverage and method of calculation. These may not be the same in both countries. The two countries may not include the same types of commodities in calculating the index numbers. Such difficulties make the index numbers only a rough guide for measuring the price levels and thus fail to give the correct purchasing power parity between the two countries.

2. Comparison of General Price Level a Difficult Problem. According to the theory, the purchasing power parity between two countries is determined by comparing their general price levels. But the price level may be made up of internally traded plus internationally traded goods, or of the internationally traded goods. If the price level is calculated in terms

of the internally traded goods, then the prices tend to equality in both countries, even allowing for the cost of transportation, tariffs, etc. Thus, according to Keynes, "confined to internationally traded commodities, the purchasing power parity becomes an empty truism."[3] On the other hand, if the price level includes both internally and internationally traded goods, then price of internally traded goods may move in the opposite direction of internationally traded goods, at least in the short period. Thus the real exchange rate may not conform to the parities.

Further, if the price level includes both types of goods, there is technical difficulty of people spending their money differently in the two countries, so that the basis for complete and accurate comparisons of price levels is lacking.

3. Not Applicable to Capital Account. Another weakness of the purchasing power parity theory is that it applies to countries whose balance of payments is determined by the merchandise trade account. It is, therefore, not applicable to such countries whose exchange rate is influenced more by capital account.

4. Difficult to Find Base Year. The theory assumes the balance of payments to be in equilibrium in the base period for the determination of the new equilibrium exchange rate. This is a serious defect, because it is difficult to find the base year when the exchange rate was initially in equilibrium.

5. Structural Changes in Factors. The theory is also based on the assumption that there have been no structural changes in the factors underlying the equilibrium in the base period. Such factors are changes in technology, resources, tastes, etc. This assumption is highly unrealistic because changes are bound to take place in these factors which, in turn, are likely to affect exchange rate.

---

3. J.M. Keynes, *A Tract on Monetary Reform*, p. 101.

6. Capital is Mobile. The theory is based on the assumption of zero-capital movements. There are many items in the balance of payments such

as insurance, shipping, and banking transactions, capital movements, etc. which are not affected by changes in the general price level. But these items affect the exchange rate by influencing the demand for and supply of foreign currencies. The theory is thus weak for it neglects the influence of these factors in determining the exchange rate.

7. Changes in Exchange Rates affect Price Level. The theory further assumes that changes in the price level bring about changes in exchange rates. But changes in exchange rates do affect the price level. For instance, if the external value of rupee falls, imports will become dearer. As a result, the costs and prices of goods using imported materials will rise. On the other hand, exports will become cheaper with fall in the external value of the rupee. Consequently, their demand will increase which will raise the demand for factors used for producing exports, and their prices will also rise. Thus changes in exchange rate do influence the price level.

8. Barter Terms of Trade Change. The theory assumes that the barter terms of trade do not change between the two trading countries. This assumption is unrealistic because the barter terms of trade constantly change due to changes in the demand for foreign goods, in the volume of external loans, in the supply of exported goods, in transport costs, etc.

9. No Free Trade. The theory is based on the assumption of free trade and *laissez-faire* policy. But governments do not follow these policies these days. Rather, they impose a number of restrictions on the movement of goods between countries. Such trade restrictions are tariff, import quotas, customs duties and various exchange control devices which tend to reduce the volume of imports. These, in turn, cause wide deviations between the actual exchange rate and the exchange rate set by the purchasing power parity.

10. Only Purchasing Power Parity does not Determine Exchange Rate. The equilibrium exchange rate may not be determined by the purchasing power parity between the two countries. Rather, a sudden increase in the demand for goods of one country may raise the demand for its currency on the part of the other country. This will lead to a rise in the exchange rate.

11. Neglect of Elasticities of Reciprocal Demand. According to Keynes, one of the serious defects of this theory is that it fails to consider the elasticities of reciprocal demand. In fact, the exchange rate is determined not only by changes in relative prices, but also by the elasticities of reciprocal demand between the two trading countries.

12. It is One Sided. Ragner Nurkse points out that the theory is one-sided in that it is based exclusively on changes in relative prices and neglects all factors that influence the demand for foreign exchange. The theory treats demand as a function of price but neglects the influence of aggregate income and expenditure on the volume and value of foreign trade, these are important factors which affect the exchange rate of a country.

13. No Direct Relation between Exchange Rate and Purchasing Power. The theory assumes direct relation between exchange rate and purchasing powers of two currencies. In reality, there is no such relation between the two.

14. Static Theory. This is a static theory because it assumes no changes in tastes, incomes, technology, tariffs, etc. These make the theory unrealistic.

15. Long Run Theory. This theory is applicable in the long run and fails to determine exchange rate in the short run.

16. Relevant for Bilateral Trade. The theory is relevant only for bilateral exchange rate determination and fails to determine exchange rate in the present multilateral trade relations.

17. Not Possible to Compute Equilibrium Exchange Rate. According to Halm[4], "Purchasing power parities cannot be used to compute equilibrium exchange rates or to gauge with precision deviations from international payments equilibrium."[4]

Conclusion. Despite these criticisms, Haberler finds the theory useful. According to him, "While the price levels of different countries may diverge, their price systems are nevertheless interrelated and interdependent, although the relation need not be that of equality. Moreover, supporters of the theory are quite right in contending that the

exchanges can always be established at any desired level of appropriate changes in the volume of money."[5]

3. The Balance of Payments Theory

According to this theory, under free exchange rates, the exchange rate of the currency of a country depends upon its balance of payments. A favourable balance of payments raises the exchange rate, while an unfavourable balance of payments reduces the exchange rate. Thus the theory implies that the exchange rate is determined by the demand for the supply of foreign exchange.

The demand for foreign exchange arises from the debit side of the balance of payments. It is equal to the value of payments made to the foreign country for goods and services purchased from it plus loans and investments made abroad. The supply of foreign exchange arises from the credit side of the balance of payments. It equals all payments made by the foreign country to our country for goods and services purchased from us plus loans disbursed and investments made in this country. The balance of payments balances if debits and credits are equal. If debits exceed credits, the balance of payments is unfavourable. On the contrary, if credits exceed debits it is favourable. When the balance of payments is unfavourable, it means that the demand for foreign currency is more than its supply. This causes the external value of the domestic currency to fall in relation to the foreign currency. Consequently, the exchange rate falls. On the other hand, in case the balance of payments is favourable, the demand for foreign currency is less than its supply at a given exchange rate. This causes the external value of the domestic currency to rise in relation to the foreign currency. Consequently, the exchange rate rises.

When the exchange rate falls below the equilibrium exchange rate in a situation of adverse balance of payments, exports increase and the adverse balance of payments is eliminated, and the equilibrium exchange rate is re-established. On the other hand, when under a favourable balance of payment situation, the exchange rate rises above the equilibrium exchange rate, exports decline, the favourable balance of payments disappears and the equilibrium exchange rate is re-established. Thus at any point of time, the rate of exchange is determined by the demand for and the supply of

foreign exchange as represented by the debit and credit side of the balance of payments. “Any change in the conditions of demand or of supply reflects itself in a change in the exchange rate, and at the ruling rate the balance of payments balances from day to day or from moment to moment.”

---

4. G.N. Halm, *Monetary Theory*, p. 228.

5. G. Haberler, *op. cit.*, p. 38.

The determination of exchange rate under the balance of payments theory is illustrated in Figure 5. *DD* is the demand curve for foreign currency. It slopes downward to the left because when the rate of exchange rises, the demand for foreign currency falls, and vice versa. *SS* is the supply curve of foreign exchange which slopes upwards from left to right. This is because when the exchange rate falls, the amount of foreign currency offered for sale will be less, and *vice versa*. The two curves intersect at *E* where *OR* equilibrium exchange rate is determined. At this rate, the quantity of foreign exchange demanded and supplied equals *OQ*. *E* is also the point where the balance of payments is in equilibrium. Any exchange rate above or below *OR* will mean disequilibrium in the balance of payments. Suppose the exchange rate rises to OR_1. The demand for foreign exchange R_1A is less than its supply R_1B ($R_1A < R_1B$). It means that there is a favourable balance of payments. When the exchange rate is more than the equilibrium rate, exports decline and imports increase. Consequently, the demand for foreign exchange will rise and the supply will fall. Ultimately, the equilibrium exchange rate *OR* will be restored where demand and supply of foreign exchange equals at point *E*. In the opposite case, when the

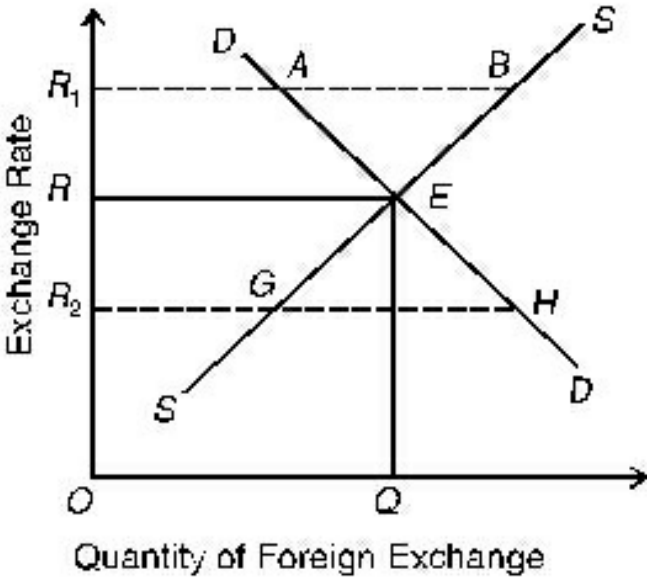


FIG. 5

exchange rate falls below the equilibrium rate to OR_2, the demand for foreign exchange R_2H is greater than its supply R_2G $(R_2H > R_2G)$. It implies an unfavourable balance of payments. But fall in the exchange rate leads to increase in exports and decline in imports. As a result, the demand for foreign currency starts falling and the supply starts rising till the equilibrium exchange rate *OR* is re-established with the equality of demand and supply of foreign exchange at point *E*.

However, it is the shape of the demand and supply curves of foreign exchange that determine the exchange rate. For this purpose, four elasticities are relevant : (*i*) the foreign elasticity of demand for exports, (*ii*) the domestic elasticity of supply for exports, (*iii*) the domestic elasticity of demand for imports, and (*iv*) the foreign elasticity of supply for imports. The equilibrium exchange rate tends to be stable if the demand elasticities are high and the supply elasticities are low.

However, according to this theory, the demand and supply of foreign exchange are determined by factors that are independent of changes in the exchange rate. Such factors are interest on foreign loans, reparation payments, etc. Further, the demand for many items that enter into import trade is perfectly inelastic so that exchange rate changes do not affect them at all. Raw materials come in this category which are required to be imported from certain countries whatever be their prices.

Criticisms. The balance of payments theory has been criticised by economists on the following counts :

1. Balance of Payments Independent of Exchange Rate. The main defect of the theory is that the balance of payments is independent of the exchange rate. In other words, the theory states that the balance of payments determines the exchange rate. This is not wholly true because it is changes in the exchange rate that bring about equilibrium in the balance of payments.

2. Neglects the Role of Price Level. The theory neglects the role of the price level in influencing the balance of payments of a country and hence its exchange rate. But the fact is that price changes do affect the balance of

payments and the exchange rates between countries.

3. No Free Tade and Perfect Competition. The theory is based on assumptions of free trade and perfect competition . This is unrealistic because free trade is not practised these days. Governments impose a number of restrictions to reduce imports and adopt measures to encourage exports. This is how they try to correct disequilibrium in the balance of payments.

4. Truism. The theory presupposes that there is an equilibrium exchange rate where balance of payments balances. This is a truism. But the equilibrium exchange rate may not be one of balance of payments equilibrium. In fact, exchange rates between countries continue to prevail under conditions of surplus or deficit in the balance of payments and there is no tendency for the balance of payments to be in equilibrium over the long run.

5. Demand for Imported Raw Materials not Inelastic. The theory has been criticised for the assumption that the demand for imported raw materials is inelastic. There is no raw materials in the world the demand for which is perfectly inelastic.

Its Superiority. Despite these criticisms, the balance of payments theory is the most satisfactory explanation of the determination of exchange rate. It studies the problem of determination of exchange rate under the framework of the general equilibrium analysis in terms of demand and supply. It studies the actual forces which lie behind the demand and supply of foreign exchange, such as the current account and the capital account of the balance of payments. An important implication of the theory is that adjustments in balance of payments can be made through devaluation and revaluation of some currency in case of deficit and surplus in balance of payments respectively. That is why, it is regarded superior to the mint par and purchasing power parity theories of exchange rate.

CAUSES OF CHANGES IN THE EXCHANGE RATE

The exchange rate between countries changes due to changes in demand or supply in the foreign exchange market. The factors which cause changes in demand and supply are discussed as under :

1. Changes in Prices. It is changes in the relative price levels that cause changes in the exchange rate. Suppose the price level in Britain rises relative to the US price level. This will lead to the rise in the prices of British goods in terms of pound. British goods will become dearer in the US. This will lead to reduction in British exports to the US. So the supply of dollars to Britain will diminish. On the other hand, the American goods become cheaper in Britain and their imports into Britain increase. So the demand for dollars will increase. Thus the supply curve for dollars will shift to the left so that the exchange rate is established at a higher level from the point of view of the US. It implies appreciation of the value of the dollar and depreciation of the value of the pound.

2. Changes in Interest Rates. Changes in interest rates also lead to changes in the exchange rate. If interest rates rise in the home country, there is a large inflow of capital from foreign countries. As a result, the exchange rate of the domestic currency will appreciate, relative to the foreign currency. The opposite will be the case, if interest rates fall in the home country.

3. Changes in Exports and Imports. The demand and supply of foreign exchange is also influenced by changes in exports and imports. If exports of the country are more than imports, the demand for its currency increases so that the rate of exchange moves in its favour. Conversely, if imports are more than exports, the demand for the foreign currency increases and the rate of exchange will move against the country.

4. Capital Movements. Short-term or long-term capital movements also influence the exchange rate. Capital-flows tend to appreciate the value of the currency of the capital-importing country and depreciate the value of the currency of the capital-exporting country. The exchange rate will move in favour of the capital-importing country and against the capital-exporting country. The demand for the currency of the capital-importing country will rise and its demand curve will shift upward to the right and the exchange rate will be determined at a higher level, given the supply

curve of foreign exchange.

5. Influence of Banks. Banks also affect the exchange rate through their operation. They include the purchase and sale of bank drafts, letters of credit, arbitrage, dealing in bills of exchange, etc. These banking operations influence the demand for and supply of foreign exchange. If the commercial banks issue a large number of drafts and letter of credit on foreign banks, the demand for foreign currency rises.

6. Changes in Bank Rate. The bank rate also influences the exchange rate. If the bank rate rises relative to other countries, more funds will flow into the country from abroad to earn high interest rate. It will tend to raise the demand for the domestic currency and the exchange rate will move in favour of the country. Converse will be the case when the bank rate falls.

7. Influence of Speculation. The growth of speculative activities also influences the exchange rate. Speculation causes short-run fluctuations in the exchange rate. Uncertainty in the international money market encourages speculation in foreign exchange. If the speculators expect a fall in the value of currency in the near future, they will sell that currency and start buying the other currency they expect to appreciate in value. Consequently, the supply of the former currency will increase and its exchange rate will fall. While the demand for the other currency will rise and its exchange rate will go up.

8. Stock Exchange Influences. Stock exchange operations in foreign securities, debentures, stocks and shares, etc. exert significant influence on the exchange rate. If the stock exchanges help in the sale of securities, debentures, shares etc. to foreigners, the demand for the domestic currency will rise on the part of the foreigners and the exchange rate also tends to rise. The opposite will be the case if the foreigners purchase securities, debentures, shares, etc. through the domestic stock exchanges.

9. Structural Influences. Structural change is another important factor which influences the exchange rate of a country. Structural changes are those which bring changes in the consumer demand for commodities. They include technological changes, innovations, etc. which also affect the cost structure along with the demand for products. Such structural

changes tend to increase the foreign demand for domestic products. It implies increase in exports, greater demand for domestic currency, appreciation of its value and rise in the exchanges rate.

10. Political Conditions. Stable political and industrial conditions and peace and security in the country have a significant influence on the exchange rate. If there is political stability and the government is stable, strong and efficient, foreigners will have tendency to invest their funds into the country. With the inflow of capital, the demand for domestic currency will rise and the exchange rate will move in favour of the country. On the contrary, if the government is weak, inefficient and dishonest and there is no safety to life and property, capital will flow out of the country and the exchange rate will move against the country.

11. Policies of Exchange Control and Protection. Policies of exchange control and protection discourage imports and lead to fall in the demand for foreign currency. As a result, the exchange rate of the home country appreciates in relation to the foreign country.

12. Type of Economy. If a country is developing, it needs to import large quantities of raw materials, and capital goods for its development along with capital. But its capacity to export is low. Therefore, its demand for foreign exchange is more which leads to the depreciation of its exchange rate vis-a-vis a developed country whose exchange rate appreciates.

EXERCISES

1. What is an equilibrium rate of exchange? How is it determined?
2. How is the rate of exchange determined under gold standard?
3. Explain how foreign exchange rate is determined under inconvertible paper currencies.
4. Critically examine the Purchasing Power Parity Theory of exchange rates.
5. Discuss the Balance of Payments Theory of foreign exchange rates.
6. Examine the factors influencing the foreign exchange rates.

CHAPTER 63

FOREIGN EXCHANGE RATE POLICY

INTRODUCTION

In the previous chapter we discussed the various theories relating to the determination of exchange rate under different exchange rate regimes. The present chapter discusses the exchange rate adjustment policies that have been in vogue from time to time with the establishment of the IMF. Before we discuss them, it is instructive to have a theoretical interlude relating to fixed and fluctuating exchange rates.

FIXED EXCHANGE RATES

Under fixed or pegged exchange rates all exchange transactions take place at an exchange rate that is determined by the monetary authority. It may fix the exchange rate by legislation or intervention in currency markets. It may buy or sell currencies according to the needs of the country or may take policy decision to appreciate or depreciate the national currency. The monetary authority (central

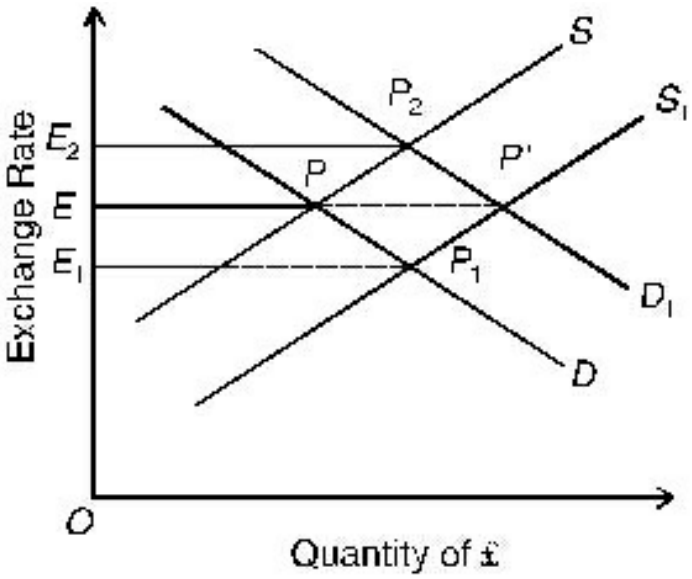


FIG. 1

bank) holds foreign currency reserves in order to intervene in the foreign exchange market, when the demand and supply of foreign exchange (say pounds) are not equal at the fixed rate. This is explained in Fig. 1 where *D* and *S* are the demand and supply curves of pound. They determine the exchange rate E which the authority maintains. Suppose the demand for pounds is more than their supply, as shown by *PP'* in the figure given the supply (*S* curve). This leads to rise in the exchange rate to E_2 when the new demand curve D_1 cuts the supply curve *S.* To maintain the exchange rate at the fixed level *E*, the monetary authority will continue to supply additional pounds to the market from its reserves till the exchange rate *E* is reached. In the opposite case, if there is excess supply of pounds equal to *PP'* in the market given the demand (*D* curve), the exchange rate falls to E_1, as shown by the rightward shift of the supply curve to S_1 and it is intersecting the *D* curve at P_1. The monetary authority will start buying these excess pounds from the market till the exchange rate *E* is reached.

The following arguments are usually advanced for and against the system of fixed exchange rates.

Case for Fixed Exchange Rates

Fixed exchange rates have the following advantages :

1. Based on Common Currency. The case for fixed exchange rate between different countries is based on the case for a common currency within a country. A country having a common currency with a fixed value facilities trade increases production and leads to faster growth of the economy. Similarly, a country would benefit if it has a fixed value of its currency in relation to other countries. Thus fixed exchange rates encourage international trade by making prices of goods involved in trade more predictable. They promote economic integration. As pointed out by Johnson, “The case for fixed rates is part of a more general argument for national economic policies conducive to international economic integration.”

2. Encourage Long Term Capital Flows. The second argument for a system of fixed exchange rates is that it encourages long-term capital

flows in an orderly and smooth manner. There is no uncertainty and risk resulting from a regime of fixed exchange rates.

3. No Fear of Currency Fluctuations. There is no fear of currency depreciation or appreciation under a system of fixed exchange rates. For instance, it removes fear that holding large quantities of foreign currency might lead to losses, if a currency's value drops. Thus it creates confidence in the strength of the domestic currency.

4. No Adverse Effect of Speculation. There is no fear of any adverse effect of speculation on the exchange rate, as speculative activities are controlled and prevented by the monetary authorities under a regime of fixed exchange rates.

5. Disciplinary. Another advantage claimed by a system of fixed exchange rates is that it serves as an 'anchor' and imposes a discipline on monetary authorities to follow responsible financial policies with countries. "Inflation will cause balance of payments deficits and reserve loss. Hence the authorities will have to take counter-measures to stop inflation. Fixed exchange rates should, therefore, impose 'discipline' on governments and stop them from pursuing inflationary policies which are out of tune with the rest of the world."[1]

6. Best for Small Countries. Johnson favours fixed exchange rates in the 'banana republics' where foreign trade plays a dominant role. Flexible exchange rates in them lead to inflation and depreciation when the exchange rate falls.[2]

7. Less Inflationary. It leads to greater monetary discipline and so to less inflationary pressures.

8. Certainty. Fixed exchange rates create certainty about foreign payments among exporters and importers of goods because they know what they have to receive or pay in foreign exchange.

9. Suitable for Common Currency Areas. This system is suitable for common currency areas such as Euro, Dollar, etc. where fixed exchange rates promote growth of world trade.

10. Promotes Money and Capital Markets. It promotes the development of international money and capital markets and helps the flow of capital among nations.

11. Multilateral Trade. This system encourages multilateral trade globally among countries because countries have no fear of wide fluctuations in exchange rates.

12. International Monetary Co-operation. The system of fixed exchange rates promotes international monetary co-operation and so helps in the smooth working of the international monetary system under such institutions as IMF, World Bank, Euro-Market.

Case Against Fixed Exchange Rates

The following arguments are advanced against a system of fixed exchange rates :

1. Sacrifice of Objectives. The principle defect in the operation of a system of fixed exchange rates is the sacrifice of the objectives of full employment and stable prices at the alter of stable exchange rates. For example, balance of payments adjustment under fixed exchange rates of a surplus country can take place through a rise in prices. This is bound to impose large social costs within the country.

2. Unexpected Disturbances. Under this system, the effects of unexpected disturbances in the domestic economy are transmuted abroad. "While a country may be protected by fixed exchange rates from the full consequences of domestic disturbances and policy mistakes, it has to bear a share of the burden of the disturbances and mistakes of others. For to the extent that excess demand 'leaks out' of the country where it was originally created, it 'leaks in' (via a balance of payments surplus) to that country's trading partner."

3. Heavy Burden. Under it, large reserves of foreign currencies are required to be maintained. Countries with balance of payments deficits must have large reserves if they want to avoid devaluation. If countries wish to remain on the fixed exchange rate system, they must hold large

reserves of foreign currencies. This also imposes a heavy burden on the monetary authorities for managing foreign exchange reserves.

4. Malallocation of Resources. This system requires complicated exchange control measures which lead to malallocation of the economy's resources.

5. Complex System. This system is very complex because it requires highly skilled administrators to operate it. It is also time consuming and may lead to uncertain results. There is always the possibility of mistakes in policy formulation and implementation.

6. Comparative Advantage Unclear. Under this system, the comparative advantage of a country is not clear. For instance, the exchange rate may be so low that a product may seen very cheap to the other country. Consequently, the country may export that commodity in which it has no comparative advantage. On the contrary with a very high exchange rate, the country may possess comparative advantage in a product.

---

1. Bo Sodersten, *op. cit.,* pp. 403-4.

2. Johnson uses the term 'banana republics' for small countries of Europe like UK, France, Denmark, etc. which are dependent on foreign trade.

7. Fixed Exchange Rate not Always Possible. Another problem relates to the stability of the exchange rate. The exchange rate of a country vis-a-vis another country cannot remain fixed for sufficiently long period. Balance of payments problems and fluctuations in international commodity prices often compel countries to bring changes in exchange rates. Thus it is not possible to have rigidly fixed exchange rates.

8. Balance of Payments Disequilibrium Persists. This system fails to solve the problem of balance of payments disequilibrium. It can be tackled only temporarily because its permanent solution lies in monetary, fiscal and other measures.

9. Dependence on International Institutions. Under this system, a country mostly depends upon international institutions for borrowing and

lending foreign currencies.

10. Problems of International Liquidity. To expand its trade, a country must have adequate international liquidity. To maintain a fixed exchange rate, the country must have sufficient reserves of foreign currencies to avoid balance of payments disequilibrium. On the other hand, excessive international liquidity is also not good for the country because the resulting extra demand may lead to international inflation.

Conclusion. In fact, a regime of fixed exchange rates presupposes uniformity of domestic policy objectives and response of prices to fluctuations in demand. Such a system would undoubtedly run into severe difficulties in the present-day world. This is because there is a reluctance to be committed to the harmonisation of domestic policy objectives; prices respond only in a limited fashion of fluctuations in the pressures of demand, and elasticities of demand in international trade have in general turned out to be quite low, at least in the short run. For these reasons, a rigidly fixed exchange rate regime has never been advanced as serious possibility in any of the recent discussions of reform of the international monetary system.[3]

FLEXIBLE EXCHANGE RATES

Flexible, *floating* or *fluctuating* exchange rates are determined by market forces. The monetary authority does not intervene for the purpose of influencing the exchange rate. Under a regime of freely fluctuating exchange rates, if there is an excess supply of a currency, the value of that currency in foreign exchange markets will fall. It will lead to depreciation of the exchange rate. Consequently, equilibrium will be restored in the exchange market. On the other hand, shortage of a currency will lead to appreciation of exchange rate thereby leading to restoration of equilibrium in the exchange market. These market forces operate automatically without any intervention on the part of monetary authority.

This is illustrated in Fig. 2 where D and S are the demand and supply curves of pounds which intersect at point P and the equilibrium exchange rate E is determined. Suppose the exchange rate rises to E_2. The quantity

of pounds supplied OQ_3 is more than the quantity demanded OQ_2. When pounds are in excess supply, the price of pounds will fall in the foreign exchange market. The value of pound in terms of dollars will depreciate. Now less pounds will be supplied and more will be demanded. Ultimately, equilibrium will be re-established at the exchange rate E. On the other hand, if the exchange rate falls to E_1, the quantity of pounds demanded OQ_4 is more than the quantity supplied OQ_1. When there is a shortage of pounds in the foreign exchange market, the price of pounds will rise. The value of pound in terms of dollars will appreciate. The rise in the price of pounds will reduce demand for them and increase their supply. This process will continue till equilibrium exchange rate E is re-established at point P.

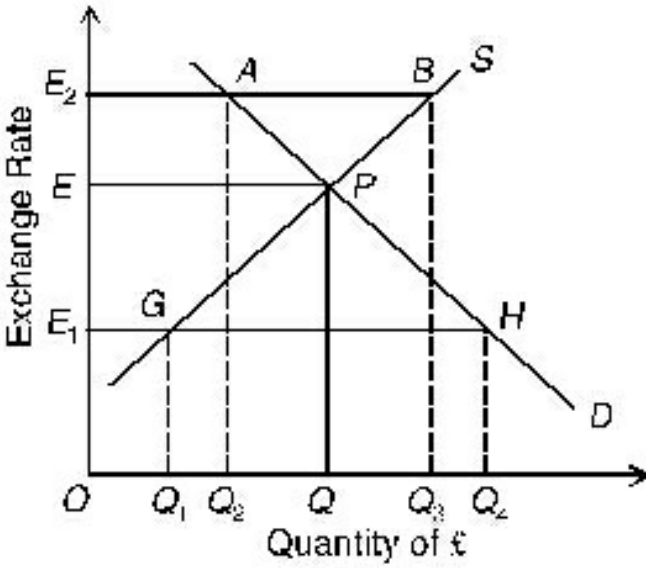


FIG. 2

---

3. Andrew Crokett, *International Money, Issues and Analysis*, 1977.

We study below the case for and against flexible exchange rates.

Case for Flexible Exchange rates

The following advantages are claimed for a system of flexible exchange rates:

1. Simple Operation. A system of flexible exchange rates is simple in the operative mechanism. The exchange rate moves automatically and freely to equate supply and demand, thereby clearing the foreign exchange market. It does not allow a deficit or surplus to build up and eliminates the problem of scarcity or surplus of any one currency. It also avoids the need to induce changes in prices and incomes to maintain or restore equilibrium in the balance of payments.

2. Smoother Adjustments. Under it, the adjustment is continual. The adjustment in the balance of payments are smoother and painless as compared with the fixed exchange rate adjustments. In fact, flexible exchange rates avoid the aggravation of pressures on the balance of payments and the periodic crises that follow disequilibrium in the balance of payments under a system of fixed exchange rates. There is an escape from the various corrective measures that are adopted by the governments whenever the exchange rate depreciates or appreciates.

3. Autonomy of Economic Policies. Under this system, autonomy of the domestic economic policies is preserved. Modern governments are committed to maintain full employment and promote stability with growth. They are not required to sacrifice these objectives of full employment and economic growth in order to remove balance of payments disequilibrium under a regime of flexible exchange rates.

4. Disequilibrium in the Balance of Payments Automatically Corrected. Since under a system of flexible exchange rates disequilibrium in the balance of payments is automatically corrected, there is no need to accommodate gold movements and capital flows in and out of countries.

5. No Need of Foreign Exchange Reserves. There is no need for foreign exchange reserves where exchange rates are moving freely. A deficit country will simply allow its currency to depreciate in relation to foreign currency instead of intervening by supplying foreign exchange reserves to the other country to maintain a stable exchange rate.

6. Removes Problem of International Liquidity. A system of flexible exchange rates removes the problem of international liquidity. The shortage of international liquidity is the result of pegged exchange rates and intervention by monetary authorities to prevent fluctuations beyond narrow limits. When exchange rates are flexible, speculators will supply foreign exchange to satisfy private liquidity needs. Individuals, traders, banks, governments and others would, of course, continue to hold liquid assets in the form of gold or foreign exchange, but these holdings would be working reserves for purposes other than the maintenance of a fixed external value of the country's currency.[4]

7. No Need of Borrowing and Lending Short-term Funds. As a corollary to the above, when foreign exchange rates move freely, there is no need to have international institutional arrangements like the IMF for borrowing the lending short-term funds to remove disequilibrium in the balance of payments.

8. Effective Monetary Policy. The system of flexible exchange rates re-inforces the effectiveness of monetary policy. If a country wants to increase output, it will lower interest rates under a regime of flexible exchange rates, the lowering of interest rates will result in an outflow of capital, a rise in the spot rate for the currency which will, in turn, cause exports to rise and imports to fall. The increased exports will tend to rise domestic prices, or income or both. Thus a favourable trade balance will reinforce the expansionary effects of lower interest rates on domestic spending, thereby making monetary policy more effective. The above process will be reversed if the country wants to fight inflation by raising interest rates. Thus a country uses monetary policy to achieve domestic objectives rather than external balance.

9. Mistakes Avoided. As a corollary , with automatic adjustments of balance of payments, there is no possibility of making monetary, fiscal and administrative policy mistakes.

10. Does not Require Complicated Trade Restrictions. A system of flexible exchange rates does not require the introduction of complicated and expansive trade restrictions and exchange controls. Thus the cost of foreign exchange restrictions is removed.

11. No Need of Forming Custom Unions and Currency Areas. Under this system, the world can get rid of competitive exchange rate depreciation and tariff warfare among nations and there shall be no need of forming custom unions and currency areas which are the concomitant results of the system of fixed exchange rates.

12. Economical. This sytem is very economical because it does not require idle holding of foreign currencies. Rather, a country can use its foreign reserves to meet its immediate requirements.

13. Promotes International Trade. This system promotes international trade because it maintains the exchange rates at their natural level through continuous market adjustments. Thus there is no danger of over-valuation or under-valuation of a country's currency.

14. Insulation from International Economic Events. Under this system, a country is protected against international economic fluctuations and shocks by making adjustments in its exchange rates.

15. Comparative Advantage. Under this system, the exchange rates are always in equilibrium. It is, therefore, possible to assess the comparative advantage of a country in a particular commodity.

Case Against Flexible Exchange Rates

The advocates of fixed exchange rates advance the following arguments against a system of flexible exchange rates :

1. Malallocation of Resources. Critics of flexible exchange rates point out that market mechanism may fail to bring about an appropriate exchange rate. The equilibrium exchange rate in the foreign exchange market at a point of time may not give correct signals to concerned parties in the country. This may lead to wrong decisions and malallocation of resources with the country.

---

4. E. Sohmen, *Flexible Exchange Rates*, 1961.

2. Official Intervention. It is difficult to define a freely flexible exchange rate. It is not possible to have an exchange rate where there is absolutely no official intervention. Government may not intervene directly in the foreign exchange market, but domestic monetary and fiscal measures do influence foreign exchange rates. For instance, if domestic saving is more than domestic investment, it means that the country is a net investor abroad. The outflow of capital will bring down the exchange rate. All this may be due to the indirect impact of government policies. Further, in the absence of any understanding among governments about exchange rate manipulation, the system of flexible exchange rates might lapse into

anarchy, for every country would try to establish favourable exchange rates with other countries. This may lead to retaliation among nations and result in war of exchange rates with disruptive effects on trade and capital movements. Thus some sort of understanding or agreement concerning exchange rates is implied in a regime of flexible exchange rates.

3. No Justification. As a corollary, there is no justification for a government to leave the determination of exchange rates to international market forces when prices, rents, wages, interest rates, etc. are often controlled by the government.

4. Exchange Risks and Uncertainty. Another disadvantage of this system is that frequent variations in exchange rates, create exchange risks, breed uncertainty and impede international trade and capital movements. For instance, an Indian who imports from Japan and promises to pay in yen runs the risk that the rupee price of yen will rise above expected levels. And the Japanese exporter who sells for rupees runs the risk that the yen price of rupees will fall below expected levels. Similarly, exchange risks may be even more serious for long-term capital movements. This is because under a system of flexible exchange rates borrowers and lenders will be discouraged to enter into long-term contacts and the possibility of varying burden for servicing and repayment may be prohibitive.

Bo Soderstен[5] has shown how flexible exchange rates increase uncertainty for traders and have a dampening effect on the volume of foreign trade. Assume that a country is under a regime of flexible exchange rates, the general price level is stable and the balance of trade is in equilibrium. Suppose the demand for the country's exports decreases, this leads to depreciation of the country's currency which, in turn, raises import prices and brings a fall in imports. Consequently, importers will be adversely affected. At the same time, exporters will gain with the increase in the prices of export goods. But the volume of exports will decline whereby they will also be losers. Opposite will be the consequences when currency appreciates. Suppose there is an abnormal inflow of short-term capital to country *A* which tends to raise its exchange rate. This will, in turn, increase the cost of *A*'s exports in terms of foreign currencies, thereby

lowering the levels of output, employment and income in its export industries. The rise of exchange rate will also lower the cost of imports, thus discouraging output and employment in *A*'s import competing industries. Thus importers and exporters will be at a disadvantage and the volume of trade will decline.

This is illustrated in terms of Sodersten's diagram, shown as Figure 3. The horizontal line *S* shows stable or fixed exchange rate, and the zig-zag line *F* shows flexible exchange rate. At time t_0 the exchange rate is the same *E*, under both flexible and fixed rate systems. At t_1 the currency depreciates and the flexible exchange rate moves to *D* while the fixed exchange rate is at the same level D_1 (= *E*). Since import prices have risen, imports will be discouraged and exports will be encouraged. At time t_2 the currency appreciates and the flexible rate moves to *A* whereas the fixed rate remains at the same level A_1 (= *E*). At *A* import prices fall. Imports are encouraged and exports are discouraged. So exports will be at a disadvantage at *A* than at A_1 and importers will gain at *A* than at A_1. Similar will be at time t_3 with fixed exchange rate at C_1 and the flexible exchange rate at *C* level. Thus fluctuations of the exchange rate around a trend value will increase risks for exports and imports that will adversely affect the volume of foreign trade.

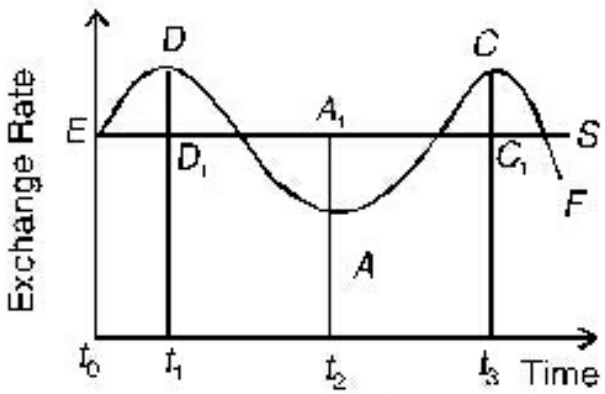


FIG. 3

---

5. Bo Sodersten, *Op. cit.*, pp. 403-4.

5. Adverse Effect of Speculation. Under this system, speculation adversely influences fluctuations in supply and demand for foreign exchange. Critics argue on the basis of empirical evidence that speculation is destabilising which means that it aggravates fluctuations in exchange rate. "It is often said that speculators see a decline in the exchange rate as a signal for further decline, and that their actions will cause the movement

in the exchange rate to be larger than it would be in the absence of speculation. In such a case, speculation is destabilising.[6] Sodersten points out that "the limited experience from the 1920s seem to show that speculation at that time was destabilising. Since floating rates became common in 1973, fluctuations in exchange rates have been large. It seems that some of the excessive fluctuations have been caused by destabilising speculation." Such fluctuations increase uncertainties in trade and reduce the volume of foreign trade further.

6. Encouragement to Inflation. This system has inflationary bias. Critics argue that under a system of flexible exchange rates, a depreciation of the exchange rate leads to a vicious circle of inflation. Depreciation leads to a rise in import prices thereby making import goods more expensive. This leads to cost-push inflation. At the same time, export prices rise. Consequently, with the rise in the cost of living, money wages rise which, in turn, intensify inflation. But an appreciation of currency is unlikely to lead to a reduction in wages and prices when imports prices fall. This is because wages and prices are sticky downwards. This leads to an asymmetry which produces that Triffin calls *ratchet effect* that imparts an inflationary bias to the economy.

7. Breaks the World Market. This system breaks up the world market. There is no one money which serves as a medium of exchange, unit of account, store of value and a standard of deferred payment. Under it, the world market for goods and capital would be divided. Resources allocation would be vastly sub-optimal. In fact, such a system clearly would not last long, according to Kindleberger.

8. Failure to Solve Balance of Payments Deficit of LDCs. LDCs are faced with the perpetual problem of deficit in their balance of payments because they import raw materials, machinery, capital equipments, etc. for their development. But their exports are limited to primary and other products which fetch low prices in world markets. Their balance of payments deficit can be removed in a system of flexible exchange rates if there is continuous depreciation of the country's currency.

---

6. C.P. Kindleberger, *International Money*, 1981.

This is illustrated in Fig. 4 where *D* is the country's demand curve for foreign exchange and *S* is the supply curve of foreign exchange. To begin, *P* is the point where *OE* exchange rate is determined. Suppose disequilibrium develops in the balance of payments of the LDC in relation to the pound currency area. This is shown by the shift in the demand curve from *D* to D_1 and the deficit equals *PP'*.

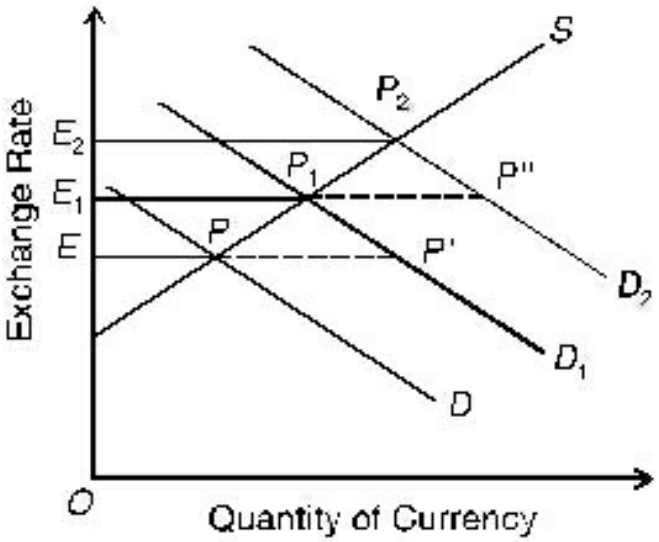


FIG. 4

This means an increase in the demand for pounds and depreciation of the currency (say, Rupee) of the LDC. Now the exchange rate of Rs. - £ rises to OE_1. This process of depreciation of the LDC currency continues with the rise in the exchange rate to OE_2 and so on. Such a policy of continuous depreciation adversely affects trade and development process in LDCs.

Conclusion. The practical use of flexible exchange rates is severely limited. Depreciation and appreciation lead to fall and rise in prices in the countries adopting them. They lead to severe depressions and inflations respectively. Further, they create insecurity and uncertainty. This is more due to speculation in foreign exchange which destabilises the economies of countries adopting flexible exchange rates. Governments, therefore, favour fixed exchange rates which require adjustments in the balance of payments by adopting policy measures.

HYBRID OR INTERMEDIATE EXCHANGE RATE SYSTEMS

There are a number of hybrid or intermediate exchange rate systems between the two extremes of rigidly fixed and freely floating (or flexible) exchange rates. They have been put into operation from time to time after the establishment of the IMF to remove the defects inherent in the two extreme systems.[7] Following a middle path, they come under *managed or*

controlled floating system. Under this system, the monetary authority (central bank) intervenes in the foreign exchange market to smooth out short-run fluctuations in exchange rates. This is done by supplying or absorbing the country's foreign exchange reserves. If the short-run demand for foreign exchange in the market is more than its supply, the monetary authority supplies the foreign exchange reserves in the market, thereby moderating devaluation of its currency. On the other hand, if the short run supply of foreign exchange is more than its demand in the foreign exchange market, it will absorb (purchase) the excess supply, increase its reserves, thereby moderating appreciation of the country's currency. This policy of managed floating is also called the *policy of leaning against the wind*.

The policy of managed floating has a number of variants.

Adjustable Peg System

It is a system in which exchange rates are pegged or fixed for a period of time. However, if a deficit or surplus of balance of payments (BOP) becomes substantial, the exchange rate is devalued or revalued. Under it, a country tries to maintain a fixed exchange rate until all its foreign exchange reserves are exhausted. If a fundamental disequilibrium in BOP still persists, the currency can be repegged at a lower or higher exchange rate. Adjusting the currency at a lower rate is called *devaluation* and at a higher rate *revaluation*. Thus under this system, exchange rate flexibility is maintained along with exchange rate stability.

---

7. For their historical evolution, refer to the last section in the present chapter.

The working of the adjustable peg system is explained in Fig. 5 where the quantity of foreign exchange (pounds) is taken along the horizontal axis and exchange rate (dollar prices of pound) on the vertical axis. D and SS' are the original demand and supply curves which determine the pegged exchange rate E at point T. If the demand for foreign exchange increases, the demand curve shifts from D to D_1 and D_2, but the exchange rate remains pegged at E. This is because the monetary authority sells its foreign exchange reserves to meet the increased demand and support this

exchange rate. The supply curve moves horizontally from point *T* to *U* and to *V* where the country exhausts all its foreign exchange reserves. But it is not possible for the monetary authority to support this pegged exchange rate *E* if the demand for foreign exchange further rises from D_2 to D_3. It will devalue its currency and repeg the exchange rate at E_1. The supply curve moves vertically from *V* to *W*. If the demand for foreign exchange continues to increase from D_3 to D_4 and D_5, the exchange rate E_1 is held on by using the country's foreign exchange reserves and the supply curve moves horizontally from point *W* to *X* and to *Z*. If the demand for foreign exchange increases further to D_6, there will be another devaluation of the currency and the exchange rate will be repegged at another higher level E_2. The supply curve will rise vertically from Z to S_1. Thus the supply curve under the adjustable peg system will be zig-zag shaped, as $STUVWXZS_1$.

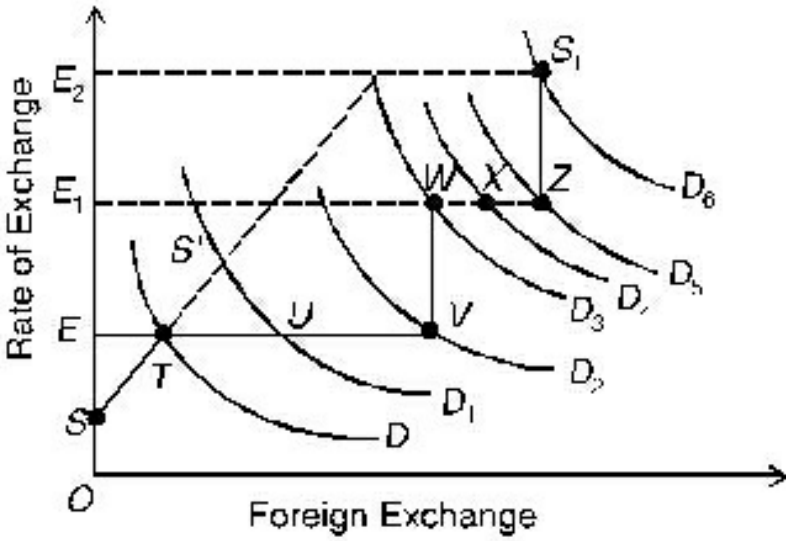


FIG. 5

Merits

The adjustable peg system helped the IMF member countries during the long boom of the 1950s and 1960s. On this basis, it has the following advantages:

1. No Uncertainty. Since rates are fixed for many years under this system, uncertainty is reduced and trade is encouraged.

2. Check of Inflation. The system of pegged rates being followed by all countries helps to bring about harmonisation of policies among them. Consequently, the world inflation remains under control.

3. Avoiding Depression and Protection. Countries can devalue their currencies in the event of a severe deficit in their balance of payments. Thus they are able to avoid depression. This also prevents them from adopting protectionist policies.

Weaknesses

This system of adjustable pegs has many flaws in its working.

1. Vague Definition. There was no operational definition of "fundamental disequilibrium." But it came to be interpreted by the IMF as meaning a persistent chronic deficit that could be ultimately corrected by devaluation.

2. Reluctance to Change Exchange Rates. There was reluctance on the part of both deficit and surplus countries to change exchange rates so as not to harm their interest.

3. Dollar Accumulation. Countries other than the United States wanted to increase their reserves which led to dollar holdings to a greater extent than needed by them. This was because dollar was a *numeraire* of the system.

4. Destabilising Speculation. This system, as it operated in the 1960s, led to destabilising speculation and consequently made controls over capital movements ineffective.

5. Expenditure-Reducing Policies. One basic flaw of this system was the insistence by the IMF on expenditure-reducing policies and the reluctancce to use expenditure switching policies for correcting disequilibria.

6. Fixity of the Peg. In the adjustable peg system, the emphasis was on the fixity of the peg rather than its adjustment.

7. Over Optimistic. Since the governments indentified deficits in balance of payments as "fundamental", they were often optimistic about balance of payments position in the future which led to the failure of this system.

Crawling Peg System

It is a system in which the monetary authority adjusts the exchange rate gradually. It adjusts the peg frequently at regular time intervals by small amounts instead of making large devaluations or revaluations when the equilibrium exchange rate changes. This is also known as *Trotting Peg* or *Gliding Parity System*.

This is explained in Fig. 6 where the *SS'* curve intersects the *D* curve at point *K* and the equilibrium exchange rate is *E*. This rate *E* is pegged between points *K* and *L* by selling foreign exchange reserves when the demand for foreign exchange increases to D_1. With a further increase in demand to D_2, the exchange rate is repegged to E_1. The monetary authority resorts to a small devaluation between point *L* and *M*. This exchange rate remains pegged at E_1 between points *M* and *N* by selling foreign exchange reserves when the demand for foreign exchange rises further to D_3. With continuous increase in the demand for foreign exchange to D_4, D_5, and D_6, the process of repegging the exchange rate at E_2 and E_3 levels and so on will continue by selling foreign exchange reserves and by small doses of devaluation at each step. Thus under the crawling peg system, we move along the supply curve in small steps, such as $KLMNPQRS_1$.

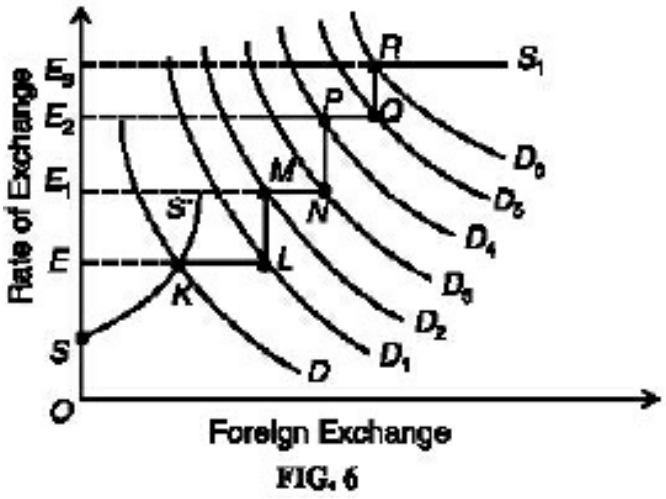


FIG. 6

This system is better than the adjustable peg system because the country resorts to small doses of inflation instead of large devaluation, as under the latter system.

Clean Float System

Under this system, the exchange rate is determined by the free market forces of demand and supply of foreign exchange. The exchange rate moves up and down freely without any intervention by the monetary authority. Thus it is simply the system of free floating exchange rate.

The working of the clean float system is explained in Fig. 7 where the original demand and supply curves D_1 and S_1 of foreign exchange (pounds) determine the equilibrium rate E at point K. When there are surpluses or deficits in balance of payments, the demand and supply curves for foreign exchange shift to the right to D_2S_2, D_3S_3 and D_4S_4, and so on. Consequently, the exchange rate moves upward along the path of points $KLMNPQR$ in case of surpluses of BOP. On the other hand, it moves along the downward path of points $KL_1MN_1PQ_1R$ in case of *BOP* deficits. In both the situations, the exchange rate fluctuates around the E level of exchange rate.

Dirty Float System

Under this system, the exchange rate is basically determined by the free market forces of demand and supply of foreign exchange. But the monetary authority intervenes from time to time to control excessive fluctuations in exchange rate. The monetary authority allows an *orderly* exchange rate adjustment when there are major changes in demand and supply of foreign exchange. But at the same time, it prevents violent fluctuations that may occur under free floating of exchange rate. The monetary authority intervenes through the sale and purchase of foreign exchange in the market. But it does allow variations in the exchange rate which moves directly from point K to M, from M to P and from P to R, as shown in Fig. 7*. As the demand for foreign exchange increases and the demand curve shifts upwards, the monetary authority sells the foreign exchange. As the supply of foreign exchange increases and the supply curve shifts to the right, the monetary authority purchases foreign exchange equal to the gap between points K and M, M and P and P and R and thus stabilises the exchange rate at E level.

Filthy Float System

When the monetary authority is compelled to intervene heavily under frequent and high fluctuations, it is called a filthy float.

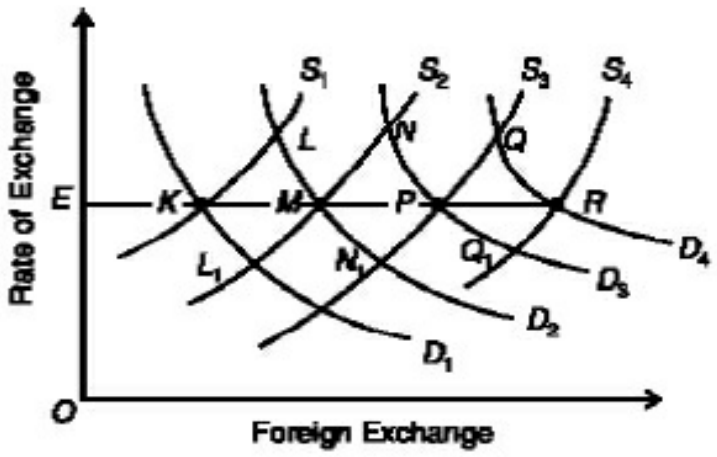


FIG. 7

Joint Float System

Under the system of joint float, a group of countries have an adjustable peg system between their *own* currencies but they have a joint float against other countries. This system is in use under the European Monetary System where its exchange rate mechanism has an adjustably pegged exchange rate band with member countries but a joint float against other currencies of the world.

Exchange Rate Band

There is another system called the exchange rate band. In this system, the currency is allowed to fluctuate between an upper and lower exchange rate about the established par value. But it is not allowed to move outside this band. Suppose the par value of exchange rate is fixed at £ 1 = $ 2 by the monetary authority. It also sets a lower limit £ 1 = $ 1.80 and an upper limit £ 1 = $2.20 to the exchange rate within which it is allowed to fluctuate freely. These limits represent the

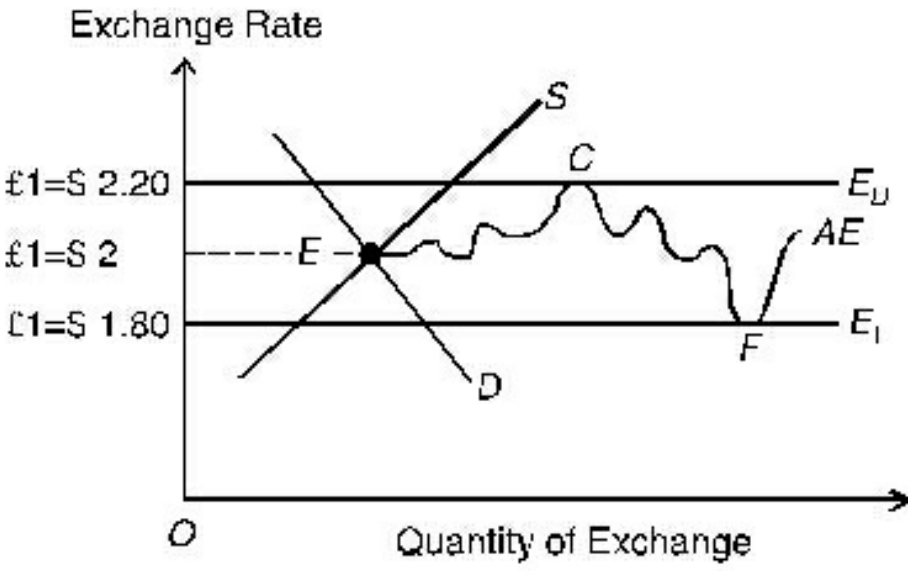


FIG. 8

exchange rate band. The exchange rate band can be narrow or wide.[8] So long as the exchange rate moves within the band, the monetary authority does not intervene. It is only when the exchange rate hits the ceiling or the floor of the band that it intervenes. This is explained in Fig. 8 where *E* is the equilibrium exchange rate (£ 1 = $ 2). The line E_U represents the upper limit of the band and E_L the lower limit. The actual exchange rate *AE* is allowed to move freely within this band. When the exchange rate reaches the upper band E_U at point *C*, the monetary authority will supply more pounds to keep the exchange rate within the band. When the exchange rate reaches the lower band E_L at point *F*, the monetary authority uses its foreign currency reserves to demand pounds to keep the exchange rate within the band.

We have explained above the exchange rate band system under the fixed exchange rate system which can also be extended to the adjustable peg and crawling peg systems.

Snake in the Tunnel

The six original *EEC* member countries started an exchange rate system similar to the joint float in 1972 which was known as *snake in the tunnel*. Under this, there was a band within the band. They allowed their currencies to fluctuate within 2.25% band relative to one-another and 4.5% relative to other currencies of the world. So fluctuations of their currencies between their band within the higher band appeared like the movement of a snake within a tunnel.[9]

MULTIPLE EXCHANGE RATES SYSTEM

It is a system under which a country adopts different rates of exchange for import and export of different commodities. A country may adopt controlled rate of exchange with some countries and free exchange rates with others. The exchange rates do not fluctuate but several fixed exchange rates and their categories may exist. Completely free and floating exchange rates may also be possible for certain transactions with some countries. The objectives of multiple exchange rates are to obtain

the maximum foreign exchange by maximising exports and minimising imports to correct the balance of payments deficit.

---

* In the case of a "note" on this system points, LL_1, NN_1 and QQ_1 are not to be shown in the figure.

8. For instance, the Smithsonian Agreement fixed the margin of fluctuations of the exchange rates to 2.25% above or below the par values of rates in 1971 and widened them to 4.5% in 1973.

Merits

The main merits of multiple exchange rates system are as follows:

1. Promotion of Exports. Even though devaluation may also be used for promoting exports but it makes imports costly and exports cheaper. But under the multiple exchange rates system, the best exchange rates can be obtained for different exports and imports. The country can obtain full advantages of elasticities of demand and supply which are favourable to it. Thus this system is more effective than devaluation.

2. Imports Profitable. A developing country has little to export but it has to import capital goods, raw-materials, technical know-how and even consumption goods on a large scale. Its imports have an inelastic demand. So it wants to enlarge the import of above goods and restricts that of luxury and other consumer goods to raise its development potential.

3. Correcting Balance of Payments Deficit. The above discussion makes it obvious that under the multiple exchange rates system maximum foreign exchange may be earned from exports and minimum possible payments can be made for import. Thus the balance of payments deficit can be corrected.

4. For Particular Country. Specially in a situation where a country has a balance of payments surplus but deficit with a particular country, the deficit can be controlled by lowering the exchange rate of commodities exported to and imported from that particular country. Thus the problem of deficit or surplus in balance of payments can be solved through the multiple exchange rates system.

5. Capital Formation. Capital goods and necessary inputs can be imported at cheaper rates through the system of multiple exchange rates. On the other hand, high export earnings may also be utilised for capital formation.

6. Capital Flows. Multiple exchange rates may be very helpful to achieve a higher capital inflow from one country and to restrict capital outflow to another country. In such a case, a higher rate of exchange would be applicable to the former and a lower exchange rate to the latter. It may also be used for channelising foreign capital into favourable lines of production.

7. Helpful for Weak Industries. Weak or declining industries can be lifted with the help of multiple exchange rates system. They may get protection or export subsidies and import capital goods, raw materials and technical know-how at some preferential and favourable rates of exchange under this system.

8. Diversifying the Economy. The system encourages the diversification of industries through favourable exchange rates. It provides protection to weak industries from foreign competition. It can help in developing new export goods industries, processing and defence industries. Commodities of mass consumption can also be produced. Thus it can diversify the economy and raise output, employment and income in the economy.

9. Maximising Revenues. The multiple exchange rates system enables the government to earn more revenues. Since this system encourages the expansion and diversification of industries, and increases output, employment and income, the government earns larger revenues from excise duties, sales tax, corporation tax, personal taxes, etc.

10. Favourable Terms of Trade. A country may secure more favourable terms of trade under this sytem. That is why it can be used for keeping prices of export goods at a higher level and prices of import articles at a lower level.

---

9. If in Fig. 8 one straight line above E_U and another below E_L is drawn at equi-distance, the figure

will depict this system.

11. Improvement of Standard of Living. Since the import of capital goods, raw materials, etc. can be obtained at low prices under this system, their cost of production is low. Similarly, essential consumer goods for mass consumption are imported cheap. These tend to reduce the cost of living and raise the standard of living of the people.

Demerits

The multiple exchange rates system has the following demerits:

1. Administrative Difficulties. In this sytem, large number of different exchange rates exist for variety of goods and for different countries. To administer them requires large administrative machinery. This involves a complex exchange control system which leads to administrative inefficiencies, red-tapism and corruption.

2. Discriminatory. This system is discriminatory because it discriminates between commodities, industries, sectors, regions and countries. The same commodity may be exported to a country at a different rate than to another country. This is likely to lead to retaliation by the other country and so adversely affect their trade and political relations.

3. Harmful for Domestic Industries. A country may import commodities at cheap rates from abroad which may harm the domestic industries as they cannot face foreign competition.

4. Not Helpful for Exports. If the demand for exports is elastic or if exports have inelastic supply in the domestic market or if foreign importers form a monopsony or oligopoly, multiple exchange rates will be of no help to the country in increasing its exports.

5. Black Marketing. This system leads to black marketing of foreign exchange. Importers buy foreign exchange at lower rates because the exchange rate for essential imports is low. But they sell foreign exchange at high rates in the foreign exchange market.

6. Limits to Different Rates. It is not possible for the monetary authority

to fix different exchange rates for a large number of exportable and importable commodities. So they are classified in small categories or groups. Their classification may be arbitrary and lead to corruption for every exporter or importer would like to have his commodity in the favourable exchange rate category.

7. Less Effective in BOP. The multiple exchange rates system is less effective than quantitative restrictions like export and import licenses, exchange controls, etc. in reducing BOP deficit.

8. Accumulation of Inventories. This system leads to accumulation of inventories. When exporters do not export their goods in anticipation of more favourable exchange rates being announced by the monetary authority, it leads to stock-piling of goods in godowns. This adversely affects production and leads to losses.

9. Not a Sufficient System. The multiple exchange rates system is not sufficient for economic development of LDCs. In such countries, the demand for essential imports is inelastic so the exchange rates for their imports cannot be lowered. On the other hand, their capacity to export is limited. So they cannot increase the exchange rates for their exports. That is why such countries suffer from shortage of foreign exchange.

Conclusion. On the whole, the multiple exchange rates system leads to malallocation of resources, reduces economic efficiency and gains from trade of the country adopting it. That is why, this system is no longer in use.

EXCHANGE RATE REGIMES IN PRACTICE

During the period preceding World War I almost all the major national currencies were on a system of fixed exchange rates under the international gold standard. This sytem had to be abandoned during World War I but most nations returned to pegged rates during the 1920s.

Adjustable Peg. With the establishment of the IMF under the Bretton Woods Agreement after World War II, exchange rates between countries

were set or pegged in terms of gold or the US dollar at $ 35 per ounce of gold. This related to a fixed exchange rate regime with changes in the exchange rate within a band or range from 1 per cent above to 1 per cent below the par value. But this was only allowed when the country could convince the IMF authorities that there was “fundamental disequilibrium” in its balance of payments. Member countries were forbidden to impose restrictions on payments and trade, except for a transitional period. They were allowed to hold their monetary reserves partly in gold and partly in dollars and sterling. These reserves were meant to incur temporary deficits by member countries while keeping their exchange rates stable. This system of adjustable pegs had many flaws in its working which led to the collapse of the Bretton Woods system in 1971 when the US Treasury refused to convert short-time liabilities into gold and made the dollar inconvertible.

Managed Floating. The sytem of adjustable peg continued till 14 August, 1971. Between 15 August, 1971 and the Smithsonian Agreement of 18 December, 1971, 48 countries including the United States, Japan and a large number of European countries abandoned fixed exchange rates. The ‘Group of Ten’ industrial countries met at the Smithsonian Institution in Washington on 18-19 December, 1971 and agreed to a new system of stable exchange rates with *wider bands*. As a first step towards realignment of major currencies, the US devalued the dollar by 8 per cent, the Japanese revalued the yen by 17 per cent and the Germans their mark by 14 per cent. The Smithsonian Agreement widened the margin of fluctuations of the exchange rates to 2.25 per cent above or below the new parities of central rates. In 1973, the band was widened to 4.5 per cent.

Another development took place in Europe when the original six EEC member countries decided to limit the fluctuations of their currencies relative to each other to a smaller band in 1972. This came to be known as “*the snake in the tunnel*”. Under this arrangement, the EEC currencies were tied together and could fluctuate within narrow limits in relation to one another but could fluctuate in relation to other currencies within the limits set by the band proposal. But due to the lack of convergence between them and world oil crisis and high inflation around the world, this system collapsed.

The Smithsonian Agreement broke down following the US dollar devaluation of 12 February, 1973. At the beginning of March 1973 India, Canada, Japan, Switzerland, the UK and several smaller countries had floating exchange rates. However, the "joint float" of the EEC countries continued even after March 1973 and was now called the "*snake in the lake*", as there was no band within which the EEC currencies could fluctuate relative to other currencies. In March 1979 the European Monetary System (EMS) was formed which created the European Currency Unit (ECU) which is a "basket" currency of a unit of account consisting of the major European currencies. Each member of the EMS fixes the value of its currency in terms of the ECU which is a pegged system of exchange rates among members' currencies.

In the meantime, the Jamaica Agreement of January 1976 formalised the regime of floating exchange rates under the auspices of the IMF. A number of factors forced the majority of member countries of the IMF to float their currencies. There were large short-term capital movements and central banks failed to stop speculation in currencies during the regime of adjustable pegs. The oil crisis in 1973 and the increase in oil prices in 1974 led to the great recession of 1974-75 in the industrial countries of the world. As a result "the dollar went into a rapid decline, which, by late 1978, had such alarming proportions that the United States government finally decided on a policy of massive intervention in order to prevent a further fall in the value of the dollar and to head off a possible financial panic". At last, the system of floating exchange rates had come to stay by 1978.

By the Second Amendment of the IMF Charter in 1978, the member countries are not expected to maintain and establish par values with the dollar or SDRs or another denominator, cooperative arrangements to maintain the value of their currencies in relation to the value of currencies of other members, and other exchange arrangements of a member's choice. The Fund has no control over the exchange rate adjustment policies of the member countries to counter disorderly conditions in the foreign exchange markets. But it lays down principles for the guidance of members' exchange rate policies. In other words, it exercises international "surveillance" of exchange rate policies of its members.

However, the system of floating exchange rates is not one of free flexible exchange rates but of “managed floating”. It has rarely operated without government intervention. Periodic intervention by governments has led the system to be called a “managed” or “dirty” floating system. In 1977, when the intervention was very heavy, it was characterised as a “filthy” float. When governments do not intervene, it is a “clean” float. But the possibilities of a clean float are very remote. Thus a system of managed floating exchange rates is evolving where the central banks are trying to control fluctuations of exchange rates around some “normal” rates even though the Second Amendment of the Fund makes no mention of normal rates.

The international experience with floating rates is difficult to evaluate, because it has been punctuated by large shocks, high inflation rates and deep recessions. Even the efforts of governments to limit exchange rate fluctuations have been marked by disagreements and have not been successful.

EXERCISES

1. Give arguments for and against a system of fixed exchange rates.
2. Do you favour flexible or fixed exchange rates? Give reasons in support of your answer.
3. What are multiple exchange rates? Give their merits and demerits.
4. Distinguish between : adjustable peg and crawling peg; clean float and dirty float; joint float and snake in the tunnel systems.
5. Write notes on : exchange rate band; snake in the tunnel; dirty float; clean float.
6. Explain the working of the system of adjustable peg. What led to the adoption of the system of managed floating exchange rates?