Interdisciplinary
Systems Research

Interdisziplinäre
Systemforschung

49

Derek W. Bunn, Howard Thomas
(Editors)

Formal Methods in Policy Formulation

Springer Basel AG

Interdisciplinary Systems Research
Analysis — Modeling — Simulation

The system science has been developed from several scientific fields: control and communication theory, model theory and computer science. Nowadays it fulfills the requirements which Norbert Wiener formulated originally for cybernetics; and were not feasible at his time, because of insufficient development of computer science in the past.
Research and practical application of system science involve works of specialists of system science as well as of those from various fields of application. Up to now, the efficiency of this co-operation has been proved in many theoretical and practical works.
The series 'Interdisciplinary Systems Research' is intended to be a source of information for university students and scientists involved in theoretical and applied systems research. The reader shall be informed about the most advanced state of the art in research, application, lecturing and metatheoretical criticism in this area. It is also intended to enlarge this area by including diverse mathematical modeling procedures developed in many decades for the description and optimization of systems.
In contrast to the former tradition, which restricted the theoretical control and computer science to mathematicians, physicists and engineers, the present series emphasizes the interdisciplinarity which system science has reached until now, and which tends to expand. City and regional planners, psychologists, physiologists, economists, ecologists, food scientists, sociologists, political scientists, lawyers, pedagogues, philologists, managers, diplomats, military scientists and other specialists are increasingly confronted or even charged with problems of system science.
The ISR series will contain research reports — including PhD-theses — lecture notes, readers for lectures and proceedings of scientific symposia. The use of less expensive printing methods is provided to assure that the authors' results may be offered for discussion in the shortest time to a broad, interested community. In order to assure the reproducibility of the published results the coding lists of the used programs should be included in reports about computer simulation.
The international character of this series is intended to be accomplished by including reports in German, English and French, both from universities and research centers in the whole world. To assure this goal, the editors' board will be composed of representatives of the different countries and areas of interest.

Interdisziplinäre Systemforschung
Analyse — Formalisierung — Simulation

Die Systemwissenschaft hat sich aus der Verbindung mehrerer Wissenschaftszweige entwickelt: der Regelungs- und Steuerungstheorie, der Kommunikationswissenschaft, der Modelltheorie und der Informatik. Sie erfüllt heute das Programm, das Norbert Wiener mit seiner Definition von Kybernetik ursprünglich vorgelegt hat und dessen Durchführung zu seiner Zeit durch die noch ungenügend entwickelte Computerwissenschaft stark eingeschränkt war.
Die Forschung und die praktische Anwendung der Systemwissenschaft bezieht heute sowohl die Fachleute der Systemwissenschaft als auch die Spezialisten der Anwendungsgebiete ein. In vielen Bereichen hat sich diese Zusammenarbeit mittlerweile bewährt.
Die Reihe «Interdisziplinäre Systemforschung» setzt sich zum Ziel, dem Studenten, dem Theoretiker und dem Praktiker über den neuesten Stand aus Lehre und Forschung, aus der Anwendung und der metatheoretischen Kritik dieser Wissenschaft zu berichten.
Dieser Rahmen soll noch insofern erweitert werden, als die Reihe in ihren Publikationen die mathematischen Modellierungsverfahren mit einbezieht, die in verschiedensten Wissenschaften in vielen Jahrzehnten zur Beschreibung und Optimierung von Systemen erarbeitet wurden.
Entgegen der früheren Tradition, in der die theoretische Regelungs- und Computerwissenschaft auf den Kreis der Mathematiker, Physiker und Ingenieure beschränkt war, liegt die Betonung dieser Reihe auf der Interdisziplinarität, die die Systemwissenschaft mittlerweile erreicht hat und weiter anstrebt. Stadt- und Regionalplaner, Psychologen, Physiologen, Betriebswirte, Volkswirtschafter, Ökologen, Ernährungswissenschafter, Soziologen, Politologen, Juristen, Pädagogen, Manager, Diplomaten, Militärwissenschafter und andere Fachleute sehen sich zunehmend mit Aufgaben der Systemforschung konfrontiert oder sogar beauftragt.
Die ISR-Reihe wird Forschungsberichte — einschliesslich Dissertationen —, Vorlesungsskripten, Readers zu Vorlesungen und Tagungsberichte enthalten. Die Verwendung wenig aufwendiger Herstellungsverfahren soll dazu dienen, die Ergebnisse der Autoren in kürzester Frist einer möglichst breiten, interessierten Öffentlichkeit zur Diskussion zu stellen. Um auch die Reproduzierbarkeit der Ergebnisse zu gewährleisten, werden in Berichten über Arbeiten mit dem Computer wenn immer möglich auch die Befehlslisten im Anhang mitgedruckt.
Der internationale Charakter der Reihe soll durch die Aufnahme von Arbeiten in Deutsch, Englisch und Französisch aus Hochschulen und Forschungszentren aus aller Welt verwirklicht werden. Dafür soll eine entsprechende Zusammensetzung des Herausgebergremiums sorgen.

ISR 49

Interdisciplinary Systems Research
Interdisziplinäre Systemforschung

Derek W. Bunn
Howard Thomas
(Editors)

Formal Methods in Policy Formulation

The Application of Bayesian Decision Analysis to the Screening, Structuring, Optimisation and Implementation of Policies within Complex Organisations

1978 Springer Basel AG

Formal methods in policy formulation: the application of Bayesian decision analysis to the screening, structuring, optimisation and implementation of policies within complex organisations / Derek W. Bunn; Howard Thomas (ed.). – 1.Aufl.
(Interdisciplinary systems research; 49)
ISBN 978-3-7643-0971-8   ISBN 978-3-0348-5288-3 (eBook)
DOI 10.1007/978-3-0348-5288-3

NE: Bunn, Derek W. [Hrsg.]

ISBN 978-3-7643-0971-8

PREFACE

This collection of papers owes its origin to a recent conference on the topic of decision analysis organised by the Royal Economic Society, during the period when one of us (H.T.) served as Programme Secretary for the Society. The papers included here were selected especially for the contribution they made to the implementation of decision analytic methods in the field of policy formulation. These selected conference papers have furthermore been supplemented by several invited contributions in order to provide a more complete exposition of the overall theme. Thus, the volume now contains a set of original papers which we believe contribute significantly to the most important aspects of this topic.

The work is grouped into two parts. Part I contains a critique of analytical methods in policy formulation and defines the essential characteristics of the policy process. Although we advocate the decision analysis approach insofar as it provides what we consider to be the most acceptable paradigm for rational action, we do so only on the balance of its merits, based upon our personal experience of consulting work in the Decision Analysis Unit at the London Business School. The papars included in this first part represent, therefore, a dispassionate evaluation of the place of analytical methods in the effective formulation of policy. Part 2 looks at some of the more important aspects in the implementation of the decision analysis approach. The methodological problems of screening, probability assessment, group consensus, multiple conflicting objectives and structuring are considered in special detail. The applications described include the siting of large-scale public facilities, the setting of standards for earthquake protection, the transportation of dangerous chemicals, the evaluation of fire control services, medical diagnosis and technological assessment.

We, as editors, should like to express our particular satisfaction in compiling together the research work of such a notable group of contributors. We also wish to acknowledge the important contribution of Charles Carter, Vice-Chancellor of Lancaster University who in his term as Secretary-General of the Royal Economic Society encouraged the development of decision analysis, economic analysis and policy formulation as a valuable conference topic.

We must also acknowledge the support of the London Business School in the development of this book. Peter Moore, the Deputy Principal, has, in particular, maintained considerable interest in the field of decision analysis, and has always been a source of encouragement and support. Our secretary, Miss Gaye Gresham, however deserves the major vote of thanks for the expert help she provided in producing the final manuscript.

Derek W. Bunn
Howard Thomas
London Graduate School of Business Studies
June, 1977

CONTRIBUTOR AFFILIATIONS

J.P. Aubin
European Institute for Advanced Studies in Management,
Place Stephanie, 20, B-1050, Brussels, Belgium.

G.B. Baecher
Department of Civil Engineering,
Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139, U.S.A.

R.E. Brooks
Graduate School of Business Administration,
University of Southern California,
Los Angeles, California 90007, U.S.A.

P. Byer
Department of Civil Engineering,
University of Toronto,
Toronto, Canada.

J.G. Gros
Council on Environmental Quality,
Executive Office of the President,
722 Jackson Place, N.W., Washington, D.C. 20006, U.S.A.

A.J. Harman
The RAND Corporation,
Santa Monica, California 90406, U.S.A.

A.S. Kalelkar
Arthur D. Little, Inc.,
Cambridge, Massachusetts, 02140, U.S.A.

K. McCusker
International Institue for Applied Systems Analysis,
A-2361 Schloss Laxenburg, Austria.

B. Naslund
Stockholm School of Economics,
Stockholm, Sweden.

R. de Neufville
Department of Civil Engineering,
Massachusetts Institute of Technology,
Cambridge, Massachusetts, 02139, U.S.A.

F.L. Offensend
Stanford Research Institute,
Menlo Park, California 94025, U.S.A.

S.J. Press
Faculty of Commerce and Business Administration,
University of British Columbia,
Vancouver, British Columbia, Canada.

C.N. Smart
Stanford Research Institute,
Menlo Park,
California, 94025m U.S.A.

J. Stringer
Australian Graduate School of Management,
University of New South Wales,
Sydney, Australia, NSW 2033

R. Tremolieres.
Universite d'Aix Marseille,
France.

D. Warner North
Stanford Research Institute,
Menlo Park, California 94025, U.S.A.

A. Williams
Department of Economics,
University of York,
U.K.

S. Zionts
School of Management,
State University of New York at Buffalo,
New York, U.S.A.

CONTENTS

PART 2

The Implementation of the Decision Analysis Approach in Policy Formulation

DEREK W. BUNN
HOWARD THOMAS

THE FIELD OF POLICY ANALYSIS

This paper defines the essential characteristics of the policy process and argues that analytical methods such as those based upon decision theory must be widened in scope in order to provide worthwhile tools for policy analysis. Indeed, when policy analysis is recognised to be a combination of social as well as analytical processes, the potential implementability of a policy itself becomes an important determinant in the final choice of the most preferred option.

1. The Field of Policy Analysis
by Derek W. Bunn and Howard Thomas

1 *The Characteristics of Policy Formulation*

The main purpose of this first chapter is to provide a short introduction and preview to Part I. In addition, the opportunity will also be taken here to outline briefly our current perspective on the nature of policy formulation. To start with, however, we should try to deal with a few minor points of clarification.

In the first instance, we do not wish to impose any distinction between the terms policy formulation and policy analysis. Clearly one could try to conceptualise these two distinct stages in an overall policy-making process; policies could be firstly formulated and then analysed. It is suggested however, such a distinction would prove meaningless in practice. The formulation of policies cannot be devoid of analysis and would moreover embody the most important issues of the final evaluation. In any case, writers on this topic tend to use these terms interchangeably and it would not increase the effectiveness of policy-making in practice to add further conceptual confusion to an already highly overconceptualised topic.

Nor do we intend to define the exact nature and boundaries of the field of policy analysis. Most fields of scientific inquiry have undefinable boundaries and yet constitute valuable categories of research and practice. Thus, for example, although there is no clear boundary between physics and chemistry, this categorisation of the sciences is still useful to us. The principal matter and approach of the physicist is characteristically different from that of the chemist. Thus, it is only important for us to seek to elucidate the principal characteristics of policy analysis which serve to identify the topic as meaningful in its own right.

Bauer and Gergen (1) classify decisions into routine, tactical and strategic with the implication that many of the characteristics of policy are derived from its principal concern with strategic decisions. This transition from routine to strategic is boardly associated with increasint complexity, wider ranging effects, longer time horizons and greater political complications. They are decisions which justify considerable time and resouces in their analysis. Dror (7) also takes this position, but with the implicit suggestion that policy formulation is associated with activities at the top of an organisation. Evidently, if policy is characterised only by scale, its nature becomes somewhat relative, situational and personal. What constitutes policy for the head of a particular organisation, maybe a tactical matter for the larger organisation of which it is only a part. The fact that policy is in certain respects identified only in the mind of the decision-maker, does not in any way invalidate its distinction as a meaningful entity. Lindblom (13) emphasises the bargaining aspects of policy formulation in its dealings with the multiple interest groups.

The decision-maker in the policy context must therefore not only be able to diagnose and analyse the problem but must also be capable of synthesising the analyses and the information from other sources in arriving at his ultimate policy choice. Because he is an actor in the total policy process he must in addition possess commonsense insights about the processes of marketing, communicating and finally implementing his policies both to his subordinates and within the organisation as a whole. Once he has this social acceptance of his immediate policy objectives he can them determine the most sensible mechanism by which those policies can be revised through time as a response to changing factors of influence and market conditions and general environmental changes.

Policy formulation is thus a process which requires the decision-maker not only to have the traditional intellectual skills of problem diagnosis, analysis and synthesis but also the political capability to implement policy within the organisation and perceive necessary directions of change if political conditions are sensitive or if major environmental changes take place. Policy-making is thus a much wider discipline than the intellectual activity known as decision-making which has captured the attention of normative decision theorists (e.g. Raiffa and Schlaifer (17), Raiffa (16), Howard (8), Moore and Thomas (14)) who postulate what a rational decision-maker should do when faced with a decision problem under uncertainty and descriptive decision theorists (e.g. Bower (2), Clarkson (4)) who examine what decision-makers actually do when faced with such problems.

II *Decision Analysis and Policy*

Keeney and Raiffa (11) distinguish between formal analysis and informal synthesis in tackling decision problems. Routine decisions are often sufficiently well understood to justify only an informal synthesis of current information in making the decision. The overall greater importance and complexity of policy formulation, however, necessitates the use of some formal method of analysis.

The present methdology of decision analysis would appear to provide the most suitable basis for the development of appropriate formalisations to aid policy analysis. Although the methods of Cost-Benefit Analysis (CBA), and Cost-Effectiveness Analysis (CEA), and Planning, Programming and Budeting Systems (PPBS) have gained considerable popularity, the method of Decision Analysis has much stronger foundations in the theory of rational decision-making under uncertainty.

Despite the relative infrequency of reported applications, there is much more experience with the use of decision analytic methods in the business sector than in the public sector (see Brown (3), Moore, Thomas, Bunn and Hampton (15) and Kaufman & Thomas (10)), although applications in the latter sector have increased over recent years (see Drake, Keeney and Morse (6)).

Decision analysis, following Raiffa (16), can be applied to a decision problem under uncertainty in terms of a series of distinct stages as follows:-

i) definition of the set of alternative actions or alternative strategies
ii) specification of a utility measure for the outcomes (or consequences) of a decision problem
iii) assessment (or definition) of a probability measure on the states of nature (or events)
iv) determination of the optimal strategy in terms of the maximisation of the expected utility crierion.

Such a formal decision analysis approach can help the decision-maker choose a good course of action by providing both a framework for choice and appropriate techniques to facilitate that choice. More importantly, the framework allows the decision-maker to incorporate his subjective judgements about probabilities and values into the formal analysis. The essential value of the approach is that it forces hard thinking about the problem situation and forces realistic examination of the set of available strategies, the generation of additional alternatives and the contemplation of scenarios which anticipate future problems and perhaps areas of growth in the activities of the organisation concerned. The net effect of a good decision analysis, therefore, is that it should highlight the areas of controversy underlying the reasons why members of a decision-making group will have differences in relation to their perception of alternatives and their valuations of consequences and sources of uncertainty.

One problem with the approach is that the decision analysis model assumes the existence of a single decision-making unit with a single individual set of utility preferences, and that this preference schedule can be used to establish the order of relative attractiveness of the possible future outcomes. In practice, in many decision situations, the decision-making unit comprises a group of people who have conflicting sets of values and preferences. How is consensus between them to be achieved and what ultimately will be the criterion for judging the most preferred course of action? Can 'optimality' in the conventional management science sense be achieved when there are so many differing viewpoints to be reconciled?

As an example, with the increasing interest of decision analysts in problems in the public sector has come the recognition that such problems involve multiple conflicting objectives for the decision-making agency, have sources of uncertainty which can only be imprecisely specified, influence different groups in society in terms of the cost-benefit picture and inevitably involve a longish time horizon with effects during the whole time period. Decision analysts have slowly reacted to this type of problem situation which requires considerably more effort and time in structuring a realistic model of the decision situation and can also incorporate a treatment of such behavioural elements as the necessity to consider the processes of negotiation and bargaining which must take place amongst the members of the decision-making group before the preferred choice of option can emerge. Equally importantly, such models must be system models and not formalisations of the individual decision-maker paradigm. If this is so, the model will require information to be drawn from all the relevant experts in the problem situation. Whilst the decision-maker remains in full charge of policy choice, such delegation of responsibility allows his attention to be focussed primarily on those aspects of the preference structure which are either crucially important or cannot realistically be delegated to others.

This development of a system model which all members of the decision-making group can reasonably accept and work with is an extremely important development in the process of bringing formal decision analytic methods into much closer proximity with the realities of the 'policy-setting' situation. In large part this is because the model provides a framework within which the responsibilities and requirements for formal analysis can be divided and delegated amongst the individuals who form the decision-making group and who then jointly become responsible for the tasks of creativity, information and gathering, evaluation and assessment of uncertainties and value measures, and finally of negotiating and bargaining as a group to decide upon their preference for a particular option.

Writers on policy analysis such as Lasswell (12) and Dror (7) have in large part failed to make a contribution to the practice of large-scale decision-making (i.e. in situations of a highly unstructured form characterised by extreme vagueness and uniqueness e.g. strategy and policy formulation for a whole organisation) because they have conducted their analyses at too high a level of abstraction. As a result they have introduced intuitively appealing concepts such as mega-policy and meta-policy but have not used these to develop a methodological approach to policy analysis that can meaningfully be used as a prerequisite for policy formulation.

We believe that recent examples of the application of formal decision analytic methods to policy questions (e.g., Howard, Matheson and North (9), de Neufville and Keeney (5)) have provided sufficient evidence to suggest that such one-of-a-kind; unique strategic policy-type questions can be handled with the formal analytical tools provided by the decision analysis approach.

However, significant problems exist in adapting the formal rational choice processes offered by the decision analysis model to the realities of the policy situation which as we have seen involve the policy-maker in the task of balancing rational analysis against the pressures emanating from the social processes and environment in which any policy analysis is carried out.

III *The Adaptation of Decision Analysis to the Policy Context*

The papers by Baecher, Gros & McKusker and Williams included in this first part present a critical analysis of the current state-of-the-art in relation to the methodology of decision analysis and its relationship to policy questions. Whereas Baecher et al provide a very thorough review of the decision analysis approach with particular reference to such questions as multiple conflicting objectives and the multi-attributed utility approach, Williams, an acknowledged expert in the field of cost/benefit analysis (CBA), discusses the integration of the CBA approach within the context of the formal methodology provided by decision analysis. As such they identify many of the problems which are faced in public sector applications:- the question of who constitutes the decision-making group; the lack of initial definition of the problem; the question of how public officials evaluate trade-offs between conflicting objectives when it is clear that individuals and groups within society will be affected in different ways by the policies when they are implemented, the issue of what CBA implies in relation to the social process as opposed to the rationalisation aspects of policy analysis.

It seems clear that adaptation of the decision analysis paradigm to the policy context requires at minimum:-

i) the identification of sensible approaches by which the individuals who form the decision-making group can search for and creatively evaluate the set of policy options. This is thus the crucial problem decomposition phase in which an explicit system model must be developed and in which policy options must be sensibly screened to provide an efficient set for subsequent evaluation.
ii) the treatment of issues and problems arising from the commonly occurring situation of multiple conflicting decision objectives in the policy domain.
iii) the treatment of questions of consensus amongst the decision-making group about assessments of uncertainty and value, i.e. the inputs to the subsequent policy analysis.
iv) the consideration of procedures of negotiation and bargaining which members of a decision-making group can use in evaluating the effects of policy implementation upon the individuals or groups likely to be affected by a particular policy.
v) the development of algorithms which can facilitate the analysis of large-scale, multiple criteria problems.

Some research work has been reported on this adaptation and we have reviewed it together with some recent original contributions in the second part of this volume. One research area that has been much more neglected is the descriptive decision (or policy) analysis area. Our level of understanding of how people actually do make policy and decision analyses is very limited. We do not know except from studies such as Bower's (2) about such issues as:-

i) How organisation attention is concentrated on specific decision problems.
ii) How decision-makers search for alternative strategies in relation to the decision problem.
iii) How decision-makers look for the consequences which should be attached to the alternative strategies.
iv) Whether decision-makers compare alternatives in terms of single criterion, such as profit, or whether they recognise the existence of consequences so intangible that they try to determine the 'satisfactory' alternative over a number of decision dimensions. This is essentially the question of how decision goals are formulated in complex organisations.
v) The extent to which decisions within organisations are effectively made by a single 'individual' decision-maker or by a group of managers. If it is the latter, the processes by which consensus in decision making was arrived at is important.
vi) The existence of information systems within organisations and how they are organised to provide an information flow for decision-makers. The types of information made available. The form of information provided for planning purposes. The control systems within the firm which impinge upon the processes of decision.

Such issues need to be evaluated through wide-ranging, in-depth case studies of policy analyses within organisations. Studies of this type take much time and considerable research effort and are often of limited generalisability in terms of other organisational situations. However, we would expect that the results of such studies will indicate that search processes and information-gathering processes constitute significant parts of decision and policy-making; particularly in less well structured situations.

The insights derived from such descriptive decision-making studies should provide further sensible suggestions for general modifications to the decision analytic approach so that it can be more efficiently applied as a policy-making procedure within the widest possible range of organisations and significant decision-making situations.

References

1. *Bauer, R.A., and Gergen, K.J. (1968) The Study of Policy Formulation, Free Press, New York.*

2. *Bower, J. (1970) The Resource Allocation Process, Division of Research, Harvard.*

3. *Brown, R.V. (1970) Do Managers Find Decision Theory Useful. Harvard Business Review.*

4. *Clarkson, G.P.E. (1963) The Trust Investment Decision Process. Prentice-Hall.*

5. *de Neufville, R., and Keeney, R.L., The Use of Decision Analysis in Airport Development for Mexico City, in Drake, Keeney & Morse (6)*

6. *Drake, A.W., Keeney, R.L., and Morse, P.M. (1972) Analysis of Public Systems. MIT Press.*

7. *Dror, Y. (1971) Design for Policy Sciences, American Elsevier, New York.*

8. *Howard, R.A. (1968) The Foundations of Decision Analysis, IEEE Transactions on Systems Science and Cybernetics.*

9. *Howard, R.A., Matheson, J. and North, W. (1972) The Decision to Seed Hurricanes. Science*

10. *Kaufman, G.M., and Thomas, H. (1976) Modern Decision Analysis, Penguin.*

11. *Keeney & Raiffa (1972) A Critique of Formal Analysis in Public Decision Making. In (6).*

12. *Lasswell, H.D. (1971) A Preview of Policy Sciences. American Elsevier, New York.*

13. *Lindblom (1968) The Policy Making Process. Prentice Hall*

14. *Moore, P.G., and Thomas H. (1976) Anatomy of Decisions. Penguin.*

15. *Moore, P.G., Thomas, H., Bunn, D.W. and Hampton, J. (1976) Case Studies in Decision Analysis. Penguin.*

16. *Raiffa, H. (1968) Decision Analysis. Addison Wesley.*

17. *Raiffa, H., and Schlaifer, R.O. (1961) Applied Statistical Decision Theory. MIT Press.*

ALAN WILLIAMS

WHAT CAN COST BENEFIT ANALYSIS LEARN FROM DECISION ANALYSIS?

The paper comprising Chapter 2 of this volume is a critical economist's assessment of how far the introduction of decision analysis methods into the usual cost-benefit approach can improve public decision-making.

2. What Can Cost-Benefit Analysis Learn From Decision Analysis?
by Alan Williams

1 *Introduction*

Cost-benefit analysis typically operates by estimating the stream of benefits associated with each of some set of alternatives, estimating the corresponding streams of costs, and reducing these to some "present value" by a time-discounting procedure. For simple cases the decision rule is "do all projects with positive net present values", though in more complex situations other rules may be applied (e.g. max NPV per unit of constrained resource). As an algorithm for making public investment decisions, it thus follows the general format of financial project appraisal, but its substantive content is very different because "benefits" do not coincide with "revenues", "costs" do not coincide with "expenditure", and the discount rate used has no necessary connection with any market rate of interest. Hence most of the effort of cost-benefit analysts is devoted to identifying, and finding ways of evaluating in money terms, the benefits and costs, and to arguing about the correct conceptual basis (and appropriate numerical value) for the discount rate.

Decision theorists, on the other hand, seem to have concentrated their attention on uncertainty as a pervasive aspect of decision-making, and especially of investment decisions, where futurity adds to uncertainty. Since the size and pervasiveness of public projects seem to add still further to the amount of uncertainty present in all aspects of the decision process, it seems reasonable to hope that cost benefit analysts might find much helpful material in decision theory, both at a conceptual level (in structuring problems for analysis) and at an empirical level (in picking up useful information about the behaviour of different sorts of decision makers in the face of uncertainty).

It is with these objectives in mind I have been scanning the literature on decision theory. I must, however, enter some defensive disclaimers at the outset. Firstly, I am not very familiar with the literature in this field, so I may have missed important relevant contributions. Secondly, I am not expert in the technicalities of statistical decision theory, so I may have failed to appreciate, or even misunderstood, the significance of that (small) part of the literature which I <u>have</u> read. Thirdly, like economists in other fields, cost-benefit analysts do not all agree with each other about the ideal nature and content of their subject, I therefore speak only for myself. Had one of my fellow cost-benefit analysts been carrying out this assignment, you would certainty have had a different paper in front of you.

Be that as it may, the plan of this paper is as follows: in section II, I present tentative taxonomy of the elements with which I shall be juggling later, and in section III, I set out my primative notions as to what decision theory has to say when applied to group decisions. In section IV, I mirror these propositions and conclusions (or lack of them) with a comparable set from the literature of public finance (and welfare economics) in which cost-benefit analysis has its intellectual antecedents. This leads me, in section V, to confess what (if anything) is actually done about risk and uncertainty in cost-benefit analysis, moving shamefacedly on to consider(in sections VI and VIII) how we might go in for self-improvement by measuring risk aversion in the public sector, firstly under a "mutual insurance" or "syndicate" model of public sector investment, and secondly under the assumption that the group decision maker is some kind of arbitrator. Then I look briefly at two other sources of uncertainty in group decision making, considering their implications under the heading "conflict and confusion costs". I end, unsurprisingly, with some conclusions, which are timid and conservative, if not downright reactionary. I would not be entirely sorry if I were convinced that I am wrong.

II *Decisions and Decision Makers*

In a recent book on decision making in business and government administration[1], Ruth Mack observes that "choice by rational man is the subject addressed by statistical decision theory. It concerns choice between predelineated alternatives. Its central application is in decision situations in which information is reasonably rich and manageable so that opinion has a firm base". (pp 55-56). In order to judge the relevance of this powerful apparatus to her chosen field, she finds it useful to divide the material up in the following manner:

The decision-agent may be "rational" ("economic") man or "natural" man, the latter differing from the former in that "his perception is selective ... his aspirations are developmental ... 'he' is typically 'they' - a decision maker is usually a collective." (p 9).

The decision process is usually an ongoing deliberative-administrative one, so rather than starting from predelineated alternatives we need to distinguish 5 stages in the process: "deciding to decide, specifying alternatives, choice, effectuation, and review." (p 9). In this context more serious errors may arise over uncertainty about what is relevant (e.g. in appropriate specification of alternatives) than uncertainty about outcomes within the selected choice set.

The decision-situation needs to be tested by six criteria: how homogeneous is the decision collective; how far can it adopt rational rather than opportunistic satisficing behaviour; what knowledge does it have about process; what knowledge does it have about the values sought; is the problem in hand to stand by itself or is it one of many for which it will be held accountable only in toto; is the choice of alternatives influenced by advancive potential (ie, a desire to change the general "structure" in which the problem in hand usually arises). (pp 10-11).

Mack also distinguishes several kinds of uncertainty which can arise within this deliberative process, each of which imposes costs upon the system:

> "(1) The uncertainty discount that is inherent in the nature of knowledge - the fact that a chance of winning a reward is less valuable than the reward for sure;
> (2) the tendency for people's behaviour to be confused by the presence of uncertainty and therefore to deteriorate;
> (3) deterioration due to externalities, to inconsistency between individual and aggregate advantage". (p 90).

She then recommends the decision rule "minimise the costs of uncertainty", which she stresses does not mean minimise uncertainty itself. She also argues that failure to apply deliberately chosen decision rules sensibly leads to a conservative bias in collectives because "it is usually easier for people to let things stay as they are than to agree to institute some particular one of several possible changes". (p 126).

In the light of this it is not surprising to find that even in the business context this "collective" aspect of decision making has pervasive effects.

> "...frequently the interests that need to be reconciled reach outside of the individual decision maker ... ultimately the individual decision maker may view himself as a surrogate of the reacting group. But this pure case is rare even in the standard format of the elected legislator. Ordinarily there is a substantial processing operation whereby the surrogate must learn of, interpret, and perhaps select among the wishes of his constituency." (p 65).

---

1. Ruth P. Mack, Planning on Uncertainty, John Wiley & Sons, New York, 1971.

"In politics and business the decision agent is often actually a group of people rather than a person representing a group ... These several sorts of collectivity as Cyert and March emphasise[1], the conflicts of goals among members ... are not capable of stable or complete resolution ... It tends to increase the uncertainty with which outcomes can be predicted..." (p 66)

"A central effect of an interpersonal decision agent is a tendency to draw out the decision process ... This implies that strategies whereby the process is governed are part of the fabric of the decision itself." (p 67)

Against this background let us look in more detail at what decision theory has to offer cost-benefit analysis, on the assumption that it is an "intendedly-rational" collective decision maker facing poorly structured decisions, whom we serve.

III *Some Apparently Relevant Material from Decision Theory*

Let us assume, as a good first approximation, that all I know about decision theory is contained in Raiffa's book Decision Analysis: Introductory Lectures on Choices under Uncertainty[2], and especially Chapter 8 of that book, which is about "Risk Sharing and Group Decisions". I propose to summarise its salient points as the starting point for my argument.

In analysing choices under uncertainty we are required to assess independently the probability attaching to each outcome, and the utility of that outcome should it occur. If outcomes are expressed in terms of money values, then for decision makers with constant marginal utility of income, the expected money value, EMV (= probability x money value of outcome) is an appropriate measure of the value of the project (lottery). For decision makers who do not view each £ gained or lost as of equal utility, then we will need to calculate expected utility (probability x utility of outcome). In order to reduce this to money terms, we could elicit the sum of money which, if offered with a probability of 1, would be deemed by the decision maker to be equivalent to the uncertainty outcome. This equivalent sum could then be substituted for the uncertain outcome to calculate the certainty monetary equivalent (CME), or value of the project (lottery) to that individual, which for those with diminishing marginal utility of income will be less than EMV.

Raiffa goes on to demonstrate that it is quite possible to offer people lotteries which none of them would accept as individuals, yet which they might accept as a group, provided some appropriate sharing arrangements were worked out. (See Appendix, Fig. 1). A lottery may be divided in many different ways to make it acceptable yet some lotteries may not be acceptable no matter how they are shared. If each sharing arrangement generates a particular distribution of expected utilities, there will be an efficient frontier of such joint utilities which satisfies the Pareto condition that no individual can increase his equated utility without that of some other individual decreasing (see Appendix - Figures 2 and 3). This can be generalised to a frontier of many lotteries also appropriately partitioned (see Appendix, Figure 4).

When it comes to deciding where to be on this frontier, two contexts of choice are considered: the Bargaining situation and the Syndicate situation.

---

1. Richard M. Cyert and James C. March, A Behavioural Theory of the Firm, Prentice Hall, Englewood Cliffs, N.J. 1963.

2. This statement is deliberately ambiguous, the implied approximate equality concealing an actual "not greater than" inequality.

In the bargaining context there is a brief mention of the possibility of eliciting interdependent utility functions (called "higher stage utility analysis"), and an even more brief mention of threat potential and deadlock, and a quick sidestep into game-theoretic formulations of the problem. Two ways of looking at the problem are distinguished: firstly, how one participant works out an optimal strategy given his expectations about the behaviour of the other(s), or secondly, how an arbitrator might set about resolving the conflict. On the latter it is concluded there are many methods one can use, but no single method seems to be universally applauded.

As regards the syndicate problem, the essential difference from the bargaining situation is that the organiser of the syndicate is free to seek partners from wherever he likes in the population, by offering whatever partitions of the original lottery he judges will be just attractive enough to them to entice them to join while minimising his own loss of expected utility (or even increasing it). Thus he is not forced to agree with other previously designated individuals.[1] This leads to a suggested "market" analogue for group decision-making.

If we concentrate on the case where a designated group has to agree on a course of action, we can distinguish 3 possible sources of difference between them, concerning respectively (a) preference for outcomes; (b) probability assessment of outcomes given an underlying state; (c) the probability assessment for underlying states (e.g. what kind of situation are we likely to face when the time comes to choose the last course of action leading to an outcome). Raiffa argues that if a panel of experts is appointed, each of whom has the same information and agrees on the basic structure of the problem, but they differ from each other on probability assessments for outcomes and on the utility attaching to each, we could then reach the paradoxical result that if the decision maker's utility and probability assessments are simply derived from those of the "experts" then it is possible for the "composite" decision to differ from that indicated by the unanimous individual votes of the panel of experts! (Table 1 Appendix) Raiffa concludes from this that it is better to use "experts" as sounding boards by which the decision maker separately elucidates his own probability assessments on the one hand and his own utility assessments on the other (subsequently using these to make his decision) than to ask experts what their respective decisions would be and then deciding by some voting mechanism. (As an aside he adds "I should do so knowing full well that I might end up choosing an action which my experts would say is not as good as an available alternative. Throughout this discussion, of course, I am assuming that I do not have to worry about the viability of my organisation, its morale, and so on" op. cit. p 233).

This conclusion makes Raiffa distinctly uncomfortable, however, I quote:

> "One can argue that a group by its very existence should have a common bond of interest. If the members disagree on fundamentals (here, on probabilities and on utilities), they ought to thrash these out independently, arrive at a compromise probability distribution and a compromise utility function, and use these in the usual Bayesian manner. In other words, the group members should consider themselves as constituting a panel of experts who advise the organisational entity; they should imagine the existence of a higher decision making unit, the organisation incarnate, so to speak, and ask what it should do. Just as it made sense to give up Pareto optimality in the panel of experts, it likewise seems to make sense in the group decision problem.

---

1. The syndicate case degenerates into the bargining case where there is a unique set of individuals involved in each and every partition that would make the lottery acceptable to him.

> But now let us consider ... the side that favours Pareto optimality. Imagine that you have fully discussed the issues with the other members of your group and that you have acquired strong feelings of your own that a_2 is better than a_1. If all your fellow group members agree with you, notwithstanding these differences in the reasons why they do so, can you imagine doing otherwise than accepting a_2 over a_1? If you thought that the group for some reason might tentatively select a_1 over a_2, would it not be your responsibility as a democratic member of the group to try to undermine a_1 in favour of a_2? And how easy your job would be. 'But', the critics of the Pareto-optimality criterion would retort, "would the organisation be as well off? Wouldn't the organisation make better decisions if the responsible parties were to thrash out their fundamental disagreements and were to build upon these compromises by maximising expected utility?
>
> These issues can be dramatised as a fight between Group Bayesians and Paretians. The Group Bayesians would argue that the behavioural assumptions for individual rationality (for example, transitivity and substitutability) are equally compelling when applied to a group acting as a decision making unit. The Paretians would argue that Pareto optimality is inviolable, and therefore the behavioural assumptions for individual rationality need to be revised when they are interpreted in a group context".

This leads Raiffa to invoke bargaining between "experts" so that one will get his way on one thing he feels strongely about, and another his way on some other thing that he feels strongely about. We thus finish up in a 'log-rolling' situation which is a variant of the bargaining situation considered earlier. His uncomfortable concluding comment is:

> "I feel that for some very cohesive groups composed of well-intentioned, responsible, idealistic members, this kind of internal log-rolling is inappropriate, that somehow the group entity is more than the totality of its members." (p 237).

IV *Some Corresponsing Material from Public Finance Theory*

Most of the foregoing points have also been thrashed over in the public finance literature, though the expositional style, context and terminology differ slightly. The major difference lies in the extent to which public sector economists have been willing to attempt precise quantification of expected utilities, being far more occupied with the logically prior problems of evaluating outcomes in money terms. These elements are usually taken as the (given) starting point in decision theory, but in cost-benefit analysis (and in much private sector financial appraisal), the bulk of the work lies in getting the problem formulated and the basic data assembled. Before considering in more detail these differences in focus of attention between the two fields, let me first of all outline the similarities.

The theory of public goods is concerned with situations in which the benefits of a service, once provided, are available to all, whether they have helped to pay for it or not. This is so because public goods differ from private goods in two crucial ways: (a) they are non-exclusive (i.e. if I consume the good, it does not reduce the amount available for you), and (b) they are non-excludable (i.e. I cannot be forced to pay for each unit consumed). Thus although it is clearly possible to arrange different levels of provision, once a level has been chosen in the pure case, it is available to everybody. Hence much of the literature has been concerned with the influence of different tax-sharing formulae upon public goods provision, assuming (in the extreme) a requirement of unanimity. This is clearly analogous to the previously designated group bargaining about lottery shares.[1]

In the public finance literature there has also raged strongly a debate over whether it is possible to separate out the choice of the optimum level of public goods supply and the optimum sharing arrangement. Samuelson's much quoted contributions[2] on this subject generated the same sort of utility-possibility frontier as Raiffa used, and led us to the conclusion that Pareto optimality, as a means of separating "efficiency" from "equity" considerations, is not enough.[3] Bergson and his followers[4] would go further and argue that we not only need a "Social Welfare Function" to select the optimum point on the utility-possibility frontier, but also to over-ride the individualistic Pareto-type calculus where individuals contravene certain ethical rules to which society subscribes (in principle, at least!)[5] Others[6] have tried to find a way out via the "higher-stage utility analysis" to which Raiffa refers, but without notable success (in the sense that although useful formal taxonomic work has been done, it lacks convincing empirical content).

As regards the "arbitration" problem in game theory, this has its analogue in the search for a suitable set of weights to apply to the incidence of costs and benefits which result from the project, so as to reflect the relative social valuation of these changes in real income, (even if we had them on a CME basis for any one individual).

---

1. An excellent exposition of this problem is to be found in Hiro Shibata - "A Bargaining Model of the Pure Theory of Public Expenditure" Journal of Political Economy Vol 79 (1) Jan/Feb 1971.

2. Samuelson, P.A. "A Pure Theory of Public Expenditure" Review of Econ & Stat Vol 36 (4) November 1954. "Diagrammatic Exposition of a Theory of Public Expenditure", ibid Vol 37 (4) November 1955.

3. McGuire, Martin C and Aaron, Henry - "Efficiency and Equity on the Optimal Supply of a Public Good" Rev. of Econ. and Stat. Vol. 51 (1) February 1969.

4. Bergson, Abram - "A Reformulation of Certain Aspects of Welfare Economics" Q.J.E. Vol. 52 1937-8, also Essays in Normative Economics, Harvard Univ. Press, 1966.

   Rothenberg, Jerome - The Measurement of Social Welfare, Prentice Hall, Englewood Cliffs, 1961.

   A good survey of the issues is presented by Peter O. Steiner "The Public Sector and the Public Interest", in Robert H. Haverman and Julius Margolis (editors) Public Expenditure and Policy Analysis, Markham Publishing Co. Chicago, 1970.

   For a very pungent statement of the anti-Pareto credo see Sidney S. Alexander's "Comment" on Arrow's paper in Julius Margolis (editor) - The Analysis of Public Output, NBER, Columbia Univ. Press, New York, 1970.

5. Even E.J. Mishan, who is, on the whole a strong Pareto-man, concedes such a role, which he calls constitution-making. See his Cost-Benefit Analysis, Allen & Unwin, London 1971, especially Chapter 45.

6. e.g. H.M. Hechman & J.D. Rodgers "Pareto Optimal Redistribution" American Economic Review, LIX (4) September 1969.

Although there have been some brave attempts at this[1] I think I could truthfully mirror Raiffa's conclusions that "there are many methods one can use" but "no single way seems to be universally applauded".

The "syndicate" problem also has its analogue in the public finance literature, as the "theory of clubs".[2] This starts from the observation that many goods are neither pure-public nor pure-private, but a mixture of both. In cases where the benefits of a service accrue to a subset of the population at large (e.g. members of a particular occupational group[3] or of a particular locality[4]) then it is possible to imagine competition (free or otherwise) between "clubs" for "members", which is the essence of the "market" aspect of syndicate problems. Arguments for greater variation between local governments in their tax-expenditure patterns are sometimes based on this analogue, which presupposes costless geographical mobility and no interjurisdictional spillovers.

Ultimately, however, there is always a hard core of group decision making which is not susceptible to market-type simulation, and where the issues and ramifications are so complex that the decision makers (politicians) call for advice from "experts" (in which class I include cost-benefit analysts). At this point I do not quite know what to make of Raiffa's oscillating state of mind. Temperamentally I prefer his "Group Bayesian" stance, since it implies that experts are not asked to make decisions, but only to elucidate the basic information on which decisions are to be made. His so-called "Paretian" stance gives "experts" a role as policy advisers, or even as a committee determining the decision. Recent experience with the Roskill Commission on the Third London Airport suggests to me that it is not a good idea to put knotty political problems "into commission" in this way, but to get analytical staff such as the Roskill Commission on the Third London Airport suggests to me that it is not a good idea to put knotty political problems "into commission" in this way, but to get analytical staff such as the Roskill Research Team to serve the government directly, i.e. to ask it to help the decision maker to clear his own mind, not to feed its results through a Commission whose recommendations the decision maker then rejects.

However, the main point I wish to pick out here from Raiffa's discussion is the clear implication that if we cost-benefit analysts do go in for utility-evaluation of a CME kind, instead of the EMV type calculations which we currently aim at, we may well confuse issues rather than clarify them, and hence maybe it would be better for us to concentrate on improving our EMV performance rather than worrying too much about expected utility. This is the key issue I want to discuss in the rest of this paper.

IV *The Treatment of Risk and Uncertainty in Cost-Benefit Analysis*

When cost-benefit analysis is used as a means of project appraisal, it tends to operate by means of calculating Net Present Values of streams of costs and benefits

---

1. See esp. B.A. Weisbrod - "Income Redistribution Effects and Benefit-Cost Analysis" in Samuel B. Chase, Jr. (editor) Problems in Public Expenditure Analysis, Brookings Institution, Washington, D.C., 1968.

2. See, for instance, J.M. Buchanan "An Economic Theory of Clubs" Economica, 1965.

   Martin C. McGuire - "Private Good Clubs and Public Good Clubs Economic Models of Group Formation" Swedish Journal of Economics 74 (1) March 1972.

3. Mancur Olson - The Logic of Collective Action, Harvard Univ. Press, 1965.

4. Charles Tiebout - "A Pure Theory of Local Expenditures" Journal of Political Economy, 1956.

at some selected discount rate. In this simple form it is open to the objection raised by Adelson[1] in relation to financial appraisal, the essence of which is as follows:

> "Since discounting, as generally defined, is truly relevant only to situations of perfect liquidity and no uncertainty, it is not surprising to find that most attempts to incorporate risk into these criteria have resulted into considerable confusion. Most writers have been satisfied to treat risk intuitively, or pretend that it does not exist. Very few have really got to grips with the problem of defining and measuring it. Thus one might allow for risk by 'shortening the expected life of the asset in the calculation' or 'estimating earnings very conservatively'. Another common suggestion (with present worth) is to use a higher discounting rate for the riskier project. How does one determine the appropriate rate for a given project? The usual answer is 'let the market decide' " (p 23).

What he suggests instead is that present worth be calculated at a risk free discount rate, using probability distributions over the set of possible outcomes, to arrive at the "conditional present worth" of a project. From this we derive an "efficient set" of project outcomes (the two significant characteristics being "expected present worth" and "variance of present worth" à la Markowitz[2]), to which he applies a (postulated) quadratic utility function to get a solution. He then concludes:

> "The decisions made by applying this approach will only be 'right' in the sense that the utility function, etc., used is right. This in turn will depend on the structure and rules of the 'economic game'. No doubt an investigation into the 'correctness' of these would be an even more valuable and rewarding study in the long run."

I do hope I do not do decision-theorists <u>as a species</u> an injustice if I regard Adelson's procedures and conclusions as characteristic. I will therefore consider his criticisms only by one to see where they bear on what we cost-benefit analysts get up to.

"Present worth" as a maximand with additive properites is certainly open to serious objections if utility is not a linear function of "present worth". But Adelson unfortunately obfuscates this issue by <u>defining</u> a capital investment problem as "one which involves resources of such magnitude that the assumption of linearity of the utility curve is not satisfactory and must be replaced". (p 36). I would define an investment problem as one in which resources are used in one time period to generate benefits in another, so that it is the <u>intertemporal</u> characteristic which is central, not mere size. Hence I would leave open for discussion the question: is this project so large that a linear utility function is a poor approximation to the <u>true</u> utility function of the decision maker? In the case of public projects subject to cost benefit analysis, they may well be large in absolute terms, and especially in their local impact, but even big projects like the Third London Airport could arguably be said to "small" in relation to total public investment so that given the margins of error elsewhere in the calculations, it is not a self-evident proposition that the assumption of a linear utility function for the <u>decision maker</u> is a serious weakness in cost benefit analysis.

---

1. R.M. Adelson - "Criteria for Capital Investment: An Approach through Decision Theory" <u>O.R. Quarterly</u>, Vol. 16 (1), 1965

2. H.M. Markowitz - <u>Portfolio Selection</u>, Wiley, New York, 1959.

There is a more subtle version of this argument, however, which turns upon some of the individual items that go into the "present worth" figures, in cases where the decision maker sees himself as a "surrogate" for a wider community. Suppose we are using measures such as "loss of market value of houses" as a means of assessing amenity losses, does not this assume that the loss of utility for each affected individual is linear in the market value of his property? This is a more serious objection, because for some people such losses will undoubtedly be large in relation to their real income, and I think that it is at this level that quadratic utility functions need to be considered more seriously. It might be countered by a dedicated CBA man that to the extent that we succeed in answering the question "what sum of money could constitute adequate compensation for the loss incurred" we are taking such non-linearity into account, and in principle this seems to me to be a complete answer. But often we do not succeed in answering that question, but are forced to resort to market-generated data from which we infer utility changes and these inferences are, I think, open to the criticism of being based on linear utility functions. (This leaves open the even more intractable question of what is implied by the way we aggregate these individuals into an overall "net" gain or loss).

On the discount rate itself, Adelson's criticisms may not be so telling in a CBA context as in a DCF context, provided one is a staunch adherent to the notation that discounting is about "time-preference", not about "opportunity cost".[1] Then the discount-rate is a statement about intertemporal utility, and although it is true, as Adelson claims, that this "weighting function (exponentially decreasing with time) ... is completely arbitrary" (p 22) so is his quadratic utility function, so we are all square at that point.

As regards the use of a "risk premium" in the discount rate as a means of allowing for risk, I would not defend such a practice in CBA, and have attacked it elsewhere on two grounds.[2]

> "In the first place, not all the phenomena associated with a particular project will exhibit the pattern of 'riskiness' which is implied by raising the discount rate, yet the discount rate will be applied to all aspects of the project, even those which can be estimated with near-certainty. A second, more pragmatic reason ... is that by appearing to deal comprehensively with risk and uncertainty ... it will also appear to make unnecessary any close analysis of the effects on the overall position of possible errors in the estimates of individual items in the project appraisal. The analysts will thereby not only have obscured important issues which ought to have been clearly identified and investigated, but they will also implicitly have arrogated to themselves the weighing of the various risks, when that judgement is essentially the task of the decision makers themselves." (para 52).

An intermediate alternative between an "expected present value" and an "expected utility" position is the pragmatic one of displaying a range of "payoffs" (in present value terms) each of which depends on a particular set of assumptions about probable outcomes, leaving the decision-maker to determine the appropriate "scenario" (and hence the optimal action, given that stance). However, if the decision is a complex one, and the decision maker is presented with an array of combinations and permutations of alternative assumptions which may make more sense to the analyst

---

1. My position as a social time preference man is clearly stated in H.G. Walsh and Alan Williams - "Current Issues in Cost-Benefit Analysis", CAS Occasional Paper No. 11, HMSO, 1969, paras 33-50.

2. See preceding reference, paras. 51 - 62, for a general statement of my views on the treatment of risk and uncertainty in CBA.

than to the decision maker, we may find ourselves in any one of the following situations:

(a) The analyst is asked to climb down off his fence, and clarify the situation by indicating quite unambiguously what his best guess is;
(b) Somebody else is asked to provide a more succinct summary of the main conclusions of the analysis, in which case he will most probably select the point estimates which he thinks are the most 'realistic'; or
(c) The decision maker selects an outcome that he likes, then reads off the set of assumptions that justify his intuitive selection.

It will then be argued that if the outcome is (a) the analyst might just as well have stuck to point estimates in the first place, that in case (b) someone else's inexpert assessment of probabilities is being substituted for that of the people who have worked on the data, and that in case (c) the analyst has merely allowed himself to become an elaborate piece of intellectual window-dressing. So all in all, it is better to keep it simple.

But before acquiescing in that weak-kneed conclusion, let us see what would be involved if we took our courage in both hands, and plunged into the problem of estimating the utility functions of public investment decision makers. This is the subject matter of the next section.

VI *Risk Aversion in the Public Sector: The Syndicate*

It is important to distinguish the two possible roles in which we might cast a public sector decision maker. He may be the organiser of a "syndicate" in Raiffa's terminology, i.e. acting as an entrepreneur for people who participate voluntarily if the terms are right, but opt out otherwise. It is this view of the situation which is considered in this section of the paper. On the other hand he may be an "arbitrator", operating on his own initiative, but dependent upon the consent of (the majority of?) the members of the group in order to stay in his job, but no individual can opt out when it comes to facing the consequences of his decisions. The implications of this view will be the subject matter of Section VII.

Most of the discussion of risk aversion in the public sector has been concerned with the former model of the situation. Thus Zeckhauser[1] argues that for goods "which by their very nature make accurate prediction of future preferences impossible" (p 98) there are at least three factors which might make it best to provide them on a collective basis:

> "(I) If the probability is small that any one individual will consume a good in a given short period of time, and if it must be kept continually available...
> (II) An essential collective-consumption element of some ... goods is that it is desirable to keep an inventory available. Thus ... plasma ... itself is of the private consumption variety, but the existence of a stockpile upon which any member can draw has a collective-consumption aspect ..."
> (III) If ... the costs of provision can be shared ... the as yet unknown consumers need not be charged any great amount. In this way some useful risk spreading can be accomplished" (p 106)

---

1. Richard Zeckhauser - "Uncertainty and the Need for Collective Action" in Robert H. Haveman and Julius M. Margolis - Public Expenditure and Policy Analysis, Markham, Chicago, 1970.

Zeckhauser further argues that competitive private insurance may only be able to cover such risks at high transaction costs, hence the government may be well placed to offer members of the community schemes of mutual insurance on better terms than any private organisation can offer.

In addition, only the government may have command of resources large enough to undertake certain risky activities, such as Research and Technological Development, major economic development projects, war or counter cyclical policies. Even here Zeckhauser sees the government in the role of "a sort of mutual investment company", and consequently he argues strongly that

> "The efficiency oriented government should evaluate the payoffs from uncertain projects in terms of the certainty equivalents of those who pay for and receive its benefits (p 112)
>
> "With this efficiency approach, it need not be the case that the government will undertake projects that yield positive expected benefits nor reject projects whose expected benefits are negative. A project that produces substantial benefits in times of general well-being will be much less attractive to individuals ... than one whose pay offs are negatively correlated with other aspects of income ..." (p 113)

This "mutual investment" cum "friendly society" view of the public sector leads to the conclusion that the risk inherent in public sector investment will be lower than that for the corresponding activity in the private sector, hence any risk-adjustment in the utility calculus will be lower.[1] Against this view we find Hirschleifer[2] arguing that if the government (virtually riskless) borrowing rate were 4 per cent, this will lead to a situation in which

> "... the marginally adopted project in the public sector would yield on average but 4 per cent while private projects with higher expected yields were failing of adoption.
>
> The opposing recommendation is based upon the contention that the higher rates required to secure funds for private investments ... are a reflection of risk aversion - and that risk aversion is a private, not a social cost" (p 270)

Against this he contends that

> "... the device of pooling provides no justification in efficiency terms for adopting what is incrementally a bad project, if in fact we can adopt the good one separately from the bad ..." (p 273)

and he would therefore limit the application of the pooling principle to those cases where

> "The key is to distinguish between private 'states' and social 'states' ... (and) ... there is only one state with respect to social totals but more than one state in terms of possible individual distributions within that total". (p 273)

---

1. See, for instance, P.A. Samuelson and W. Vickery, "Discussion", American Economic Review: Papers and Proceedings Vol. 59 (May 1964) pp 88-96.

2. J. Hirschleifer - "Investment Decisions Under Uncertainty: Applications of the State Preference Approach" Quarterly Journal of Economics Vol. 80 (May 1966) pp 252-277

Consequently, his conclusion is that for public investment in general

> "... the efficient discount-rate, assuming perfect markets, is the market rate implicit in the valuation of private assets where returns are 'comparable' to the public investment in question - where 'comparable' means having the same proportionate time-state distribution of returns." (p 276/7)

Operationally, Hirschleifer's dictum seems to imply finding an analogue in the private sector to the payoff matrix of each public investment, discovering on what terms private capital was raised for it (assuming that such capital is project-specific, and the suppliers of capital were well-informed about the project), and then using this as the discount rate in evaluating the public investment.

Arrow and Lind[1] accept Hirschleifer's general framework of analysis, but argue that

> "when the risks associated with a public investment are publicly borne, the total cost of risk-bearing is insignificant and, therefore, the government should ignore uncertainty in evaluating public investments. Similarly, the choice of the rate of discount should in this case be independent of considerations of risk. This result is obtained not because the government is able to pool investments but because the government distributes the risk associated with any investment among a large number of people. It is the risk spreading aspect of government investment that is essential to this result".

This is also the intellectual foundation of the position of H.M. Treasury in the justification for its "test discount rate" to be applied to investment decisions by nationalised industries, but their position is more guarded than Arrow's stark prescription "ignore uncertainty" for the authors of the White Paper[2] wrote:

> "The test rate of discount, being a uniform rate for all industries, does not include allowance for the risks of individual investments. Exercising judgement as to what risks are worth taking is essentially a function of management ..." (para 12).

VII *Risk Aversion in the Public Sector: The Arbitrator*

This leads us away from the "syndicate syndrome" into the realm of the "manager" or "arbitrator". As Arrow[3] observes:

> "This position rejects the notion that individual preferences as revealed by market behaviour are of normative significance for government investment decisions, and asserts that time and risk preferences relevant for government action should be established as a matter of national policy ..."

---

1. Kenneth J. Arrow & Robert C. Lind - "Uncertainty and the Evaluation of Public Investment Decisions" American Economic Review Vol. 60 (1970) pp 364-378; reprinted in Arrow's Essays in the Theory of Risk-Bearing" Markham - North Holland, 1971.

2. Nationalised Industries: A Review of Economic and Financial Objectives, Cmnd. 3437, H.M.S.O., November 1967.

3. Essays in the Theory of Risk Bearing, op cit (1) p 241

This is basically the position of Eckstein[1] and Marglin[2], and also my own.

There appear to be two alternative routes here, the first being to regard the "arbitrator" as the unit for analysis, and work on his utility function as the basic element in the situation, and the second being to imagine ourselves trying to help him elicit (via non-market mechanisms) the community's collective view on this matter.

If we took the latter path we would confront the kinds of issues recently enumerated by Dasgupta and Pearce[3], viz:

> "A more precise analysis of the concept of 'a socially appropriate degree of risk aversion' is an extremely difficult task, which we shall not attempt to undertake. However, some general comments are in order.
>
> Firstly, the appropriate degree of risk aversion must depend on the general levels of prosperity. Richer communities can afford to take more risks - for example, to undertake more research which tends to involve a very high variance of outcomes.
>
> Secondly, both for an individual and for a society, the appropriate degree of risk aversion depends on whether the existing liquidity position is regarded as satisfactory. A society with large accumulated foreign exchange reserves may, for example, have a different attitude to the riskiness of an export project rather than one with a low level of reserves and a persistent balance of payments problem ...
>
> Thirdly, the degree of risk aversion may be influenced by religious or moral values. A strong aversion to risky undertakings may for example be associated with the prevalence of a puritanical ethic.
>
> To conclude, the problem of properly assessing risk aversion is a difficult one". (pp 186-7).

In this role the analyst is mainly guide, philosopher and friend (not to mention electoral agent) and at this generalised abstract level it is a task I cheerfully leave to the political scientists and others of that ilk. There is a more useful role for quantitative analysis in working out the implications of various possible stances, and in ensuring that any policy statements which emerge are not so vacuous or ambivalent as to be useless for decentralised decision making.

The alternative approach is to try to elicit, by experiment or observation, the behaviour of the "manager" towards risk in any fairly simple (earlier) decisions he makes, so as to "advise him" what this revealed attitude to risk implies for more complex (later) decisions. This looks more hopeful from a decision theory viewpoint, but there are some snags.

---

1. Otto Eckstein - "A Survey of the Theory of Public Expenditure Criteria", in J.M. Buchanan (editor) Public Finances: Needs, Sources and Utilization, NBER, Princeton University Press, 1971. esp. pp. 468-478.

2. Stephen A. Marglin, - Public Investment Criteria, Allen and Unwin, 1967, pp. 71-74.

3. Ajit K. Dasgupta and D.W. Pearce - Cost Benefit Analysis: Theory and Practice, MacMillan, 1972.

The first batch of difficulties stem from the limitations of the decision-theory framework itself as catalogued, for instance, by Menges.[1]

> "Application of statistical decision theory to social phenomena is limited by the instability of the decision maker's targets, changes in the evaluation of outcomes, etc.
>
> Let us imagine that a 'real decision maker' ... wants to consult us, the decision-theorists. In order to help him, we have to ask him:
>
> > Can you enumerate exactly your possible actions?
> >
> > Do you know exactly the possible states of nature?
> >
> > Do you possess an unlimited sensitivity for utility differences?
> >
> > Do your utilities possess the von Neumann-Morgenstern property?
> >
> > Do you know objectively - either a priori or a posteriori - the distribtuion law over the states of nature, or, if not, are you willing to accept the pure minimax rule as the expression of rational behaviour in your actual decision situation?
> >
> > Do you know that the decision situation is stable within a certain, not too short, period?
>
> If he is able to answer these questions in the affirmative then we can apply the Bayes or minimax criterion or some combination of them.
>
> Otherwise, we can possibly give him some help, on the basis of previous expeience, to fill up the gaps in his knowledge.
>
> But what should we do if some intrinsic feature of the situation dictates a negative answer to one of the above questions?"

Menges used this questionnaire to buttress his argument that getting the decision context right (what he calls the "pre-decision decisions") is often the crucial thing, and although he probably overstates the difficulties, it would be foolish to plunge in thinking that attitudes to risk are going to be "revealed" at all readily or tidily from past decisions.

In a slightly different context Foster[2] has argued recently that the "post-mortem" approach to eliciting information about public investment decisions can run into formidable difficulties in practice. Amongst the obstacles are that

> "Hardly anybody stands to gain from backchecking ...

---

1. G. Menges - "On Some Open Questions in Statistical Decision-Theory", in Karl Borch and Jan Mossin (editors) Risk and Uncertainty, Macmillan, 1968 pp 140-154.

2. C.D. Foster - "Policy Review" a paper given to a I.M.T.A. Conference on "Programme Budgeting in 1984" Portsmouth, September, 1972.

> Quite commonly backchecking is frustrated because the data ... have disappeared; or ... those who made the calculations have either moved or are fully employed elsewhere, or themselves scratch their heads when they look at their rough notes and cannot make head or tail of them ...
>
> ... One usually finds that backchecking runs into the sands because it becomes impossible to agree on a satisfactory procedure for checking all the assumptions underlying the forecasts. But though this problem is overcome, there are two other flaws in the process. The first is that it is inevitably easier to backcheck on the chosen than on the rejected alternatives. The second is that even if lower management did get one project seriously wrong ... provided on average their projects show a satisfactory return, a high proportion of failures may be tolerable, particularly if the costs of reducing risks in relation to each single project are high ... What would be fair then would be the backchecking of virtually every project; but this would involve such an expenditure of resources that it is never contemplated ..."

This leaves us with the experimental method, and in this context presumably some kind of simulation by gaming. On this subject Edwards[1] comments as follows on various experiments designed to establish ordered metric scales of utility in gambling decisions:

> "A most disturbing possibility is raised by experiments by Marks[2] and Irwin[3] which suggest that the shape of the subjective probability function is influenced by the utilities involved in the bets. If utilities and subjective possibilities are not independent, then there is no hope of predicting risky decisions unless their law of combination is known, and it seems very difficult to design an experiment to discover that law ..." (Penguin Volume, p.40)

Since politicians (and other public sector decision makers) are no less frail and childlike than the rest of us when it comes to selective perceptions of risks, this could seriously hamper experiments in this field, as could another of Edwards' observations of people's behaviour in bargaining situations:

> "The main finding from these studies of multi-person games seems to be that people import into bargaining situations a strong desire for equity. Equity-seeking is promoted by effective and free communication and seriously hindered or even prevented by severely restricted communication. Equity-seeking produced results in conflict with those which game theory and similar theories about rapacious economic men imply, except in those games in which equity-seeking and

---

1. W. Edwards - "Theory of Decision Making" Psychol. Bull, Vol. 51 (1954) pp 380-417; and "Behavioural Decision Theory" Ann. Rev. Psychol. Vol. 12 (1961) pp 473-498 (Both reprinted in W. Edwards and A. Tversky (editors) Decision Making, Penguin, 1967).

2. Rose W. Marks - "The effect of probability, desirability, and 'privilege' on the stated expectations of children", J. Pers. Vol. 19 (1951), pp 332-351.

3. F.W. Irwin - "Stated Expectations as functions of probability and desirability of outcomes" J. Pers. Vol. 21 (1953) pp 329-335.

uninhibited rapacity have the same consequences. If this finding stands up under more experimentation, especially with larger payoffs, theories about multi-person decision situations must either be modified to incorporate such social motives explicitly, or else some means for incorporating them in utility functions must be found". (Penguin, pp 88-89).

This seems to leave us deadlocked once more, and perhaps we cannot do better than follow Marglin's counsel:

> "The main point ... is that the existence of risk requires policy makers to specify their attitudes towards risk. Specification might be explicit, for example, in terms of the rate at which they are willing to trade greater average returns for less variance or skewness in returns; or specifications might be implicit in terms of constraints on variance and skewness". (p 72)

> "Thus the basic points, which are valid for both risk and uncertainty, are, first that policy makers must specify their attitudes towards fluctuations in costs and benefits rather than abdicate this value judgement to subordinates who introduce their own biases under the guise of technical criteria akin to the safety criteria for loading bridges. Second, the public sector should take advantage of the facility that the size and variety of public sector investments offer for pooling risks and uncertainties..." (p 74).

VIII *Conflict and Confusion Costs*

At this point let us return to Mack's threefold classification of uncertainty

(1) risk aversion per se;

(2) deterioration of decision capability due to confusion;

(3) inconsistency due to conflicts of interest within the group;

Most of the preceding discussion has been about (1), with some consideration of (3) as envinced in bargaining situations. But we have not yet given any thought to (2), apart from mentioning en passant some of the short-cuts that people use in order to get their problems simplified.

An interesting experiment concerning alternative decision structures was conducted by Bower[1]. He compared the decision making behaviour of two types of groups, both of which consist of individuals who act so as to maximise expected utility. The first type of group (a "foundation") has a decision function for ordering all acts in a transitive manner satisfying the dominant principle and "corresponds to a neo-classical economy where there is a social welfare function". In the second type of group (called a "rational team") "all members have identical utility functions over certain outcomes" but "their judgemental probability distributions (jpd's) over uncertain

1. Joseph L. Bower - "The Role of Conflict in Economic Decision-Making Groups: Some Empirical Results" Quarterly J. of Econ. Vol. 79 (2) May 1965, pp. 263-277.

states differ. Faced with uncertainty they amalgamate their jpd's into a group jpd and maximise expected utility". Thus:

> "The team problem is one of reconciling conflicting jpd's. The foundation problem is a bargaining game where information and threats are used as persuasive weapons ... interpersonal conflict is absent in the team problem and present in the foundation problem. We expect teams to make superior choices as measured by the U. score ..." (p 271).
>
> "The results for the full sample confirmed the basic hypothesis as did the data for groups where unanimity was required. Teams did better on the average than foundations and unanimous teams did better than unanimous foundations. On the other hand, the performance of teams under majority rule was superior to that of foundations.
>
> Further analysis indicates that the reason for the poorer overall performance of foundations under a requirement of unanimity is that four of the foundation groups failed to reach agreement as opposed to only one team. When the sample is truncated so as to remove all cases of no choice, the average performance of foundations was superior to that of teams in every instance..." (p 272)
>
> "... the analysis suggests that a multi-activity model of group decision-making may be appropriate in which conflict plays at least a dual role... Suppose ... that there are really three activities involved ... (1) finding alternatives and sharing relevant information - we call this search (2) examining the relationships among information possessed and the relative appropriateness of defined alternatives - we call this analysis, and (3) making a decision - we call this choice. Then, in a foundation, the personal commitment of a subject to an initial position motivates him to defend his choice by presenting all the information which supports his position in as cogent a manner as possible. In other words, group search is stimulated in both extent and quality ..." (p 273)
>
> "The same conflict may motivate superior analysis. Because the member of a foundation has something to lose when he shifts his position, he has a natural incentive to evaluate carefully all information he possesses in relation to proposed alternatives. Bargaining, in short, is a kind of internal pricing system for information, aiding analysis ...
>
> Proceeding with this model of the group decision process, the data on the occurrence of no choice indicate that the probability of reaching agreement decreases as interpersonal conflict increases ..." (p 275).

This suggests that there may be an optimum level of conflict within a group ... just enough to stimulate search and analysis but not so much as to deliberate the capacity for choice. It would be interesting to discover whether getting a "foundation" (e.g. the Cabinet) to agree upon an explicit policy about "risk" (e.g. on a major issue, such as entry into the EEC, where there is considerable uncertainty about states, outcomes, their probability and time dimensions, not to mention their utility) would push them beyond this optimum or leave them within it.

IX *Conclusions*

I have a distinct impression that decision theorists have not grappled with group decision-making any more successfully than anyone else, once one gets away from the "syndicate" syndrome. Since I regard most public investments to which cost-benefit analysis is applied as being essentially matters of collective choice (i.e. the group has to make a decision and enforce it if necessary on demurring members) I think this voluntaristic analogue is severely limited in its usefulness, even though I would concede that there are areas of public sector activity in which "club membership" can be considered effectively as a matter of individual choice (e.g. by moving from a local authority with a comprehensive educational system to one without, or vice versa) within limits and often at high personal cost.

To the extent that the group decision maker is interested in eliciting the attitudes towards risk of those who are affected by his decisions, then if they satisfy von Neumann-Morgensterm axioms and suitable experiments can be devised and conducted, this would be a useful additional component in a cost benefit analysis. But this would not enable the group decision maker to avoid the necessity for deciding what weight he would give to these individualistic assessments, both in the process of amalgamating them into a "group" view, and in the process of determining what weight should be given to factors which they might not have taken into account (moral obligations to other groups, concern for posterity, the possible effect of this decision upon others, etc).

Encouraging the group decision maker to be more explicit (even if not for wide publication!) about his attitude to risk per se would undoubtedly be useful and in full conformity with the general thrust of cost benefit analysis. It may be that some limited experimental evidence could be granted here, though how far this could confidently be carried over to major issues such as British Entry into the EEC (or even the siting of the third London Airport) is open to question. Still, much cost-benefit analysis is concerned with much more hum-drum issues, like road improvements or water supplies, and this may be a better sphere in which to start.

Meanwhile, I think the important immediate lesson for us cost-benefit analysts is that we should try much harder to encourage decision makers to structure their problems in a manner more likely to render them susceptible to analysis by the sophisticated paraphernalia of Bayesian Inference, expected utility, and Markowitz portfolio selection. Identifying the core problem and structuring it for analysis is still a major undertaking, and to be able to construct a decision tree that does not develop impenetrably cancerous growth is a skill greatly to be valued. We cost-benefit analysts have a crucial role to play in identifying, measuring and valuing the elemental costs and benefits and their incidence in each eventuality. Estimating probabilities may also be our forte if they concern economic variables, but much of the uncertainty typically concerns non-economic variables (the incidence of disease, or crime, or road accidents, or rainfall, or technological advance) in which we cannot ourselves be expert. It is only after all this has been done, if I have understood matters correctly, that decision theory comes to our aid.

I shall, therefore, continue my backward practice of recommending that we stop short at "sensitivity analysis" in presenting the results of cost-benefit analyses to decision makers together with such commentary about balance of probabilities as the "experts" can provide. It is then up to the decision maker to decide which risks he will run, but he should clearly be encouraged to give his judgement in writing ... just in case we pass this way again.

ACKNOWLEDGEMENTS

This paper arose out of work on the application of cost-benefit and cost-effectiveness analysis to public sector problems in the UK, conducted in the Institute of Social and Economic Research in the University of York, and financed as part of a programme of public sector studies by the Social Science Research Council.

APPENDIX

A.1 All the material in this appendix is taken from Chapter 8 of Howard Raiffa - Decision Analysis: Introductory Lectures on Choices Under Uncertainty, Addison Wesley, 1968.

A.2 Group Acceptance and Rejection Sets

Suppose that two individuals are asked whether they are willing to accept a lottery in which there is a .5 chance of winning \$1000 and a .5 chance of losing \$500. Suppose that individual 1 will accept any lottery with these probabilities which falls on or to the SE of g_1 in figure 1 (Raiffa's Figure 8.2, p 193), while individual 2 will accept any on or to the SE of g_2. Hence neither individual would accept the lottery. If each were offered an equal share in the lottery, it becomes a (-250, +500) lottery (still with a probability of .5 for each outcome), which falls within the acceptance set delineated by g_2, but outside that of g_1. So an equiproportional sharing arrangement would not entice this "group" (of 2 people) to accept the lottery. However, if individual 2 were offered (-350, +500) he would accept it, and the residue (-150, 500) would be acceptable to individual 1, so in this case there is a set of nonproportional sharing arrangements which would make the lottery acceptable to the group.

A.3 The utility possibility frontier: one lottery shared in various ways

If the given lottery were partitioned in various ways, and the utility to each of his share plotted, we could derive an "efficient frontier" of such joint utility evaluations as in Figure 2 (Raiffa's figure 8.3 p 198) if the acceptable set includes improvements on the status quo, or as in Figure 3 (Raiffa's figure 8.4 on p 198) if no partition would make the lottery acceptable. In figure 2, g is not an "efficient", or Pareto-optimal, situation since there are other configurations, such as d, e and f, which both parties would prefer. If the efficient frontier has "dips" in it, such as the segment on which c lies, this could be filled in by randomisation, i.e. by tossing a coin to determine whether sharing arrangement b or d shall operate, which would generate an outcome h midway between them (and superior to c). It is not possible, however, to determine which of the outcomes on this frontier is best without assigning weights to the respective utilities, $\bar{u}_1$, and $\bar{u}_2$.

A.4 The utility possibility frontier: many lotteries shared in various ways

This process could be replicated for several lotteries, each having its own joint utility distribution, as in Figure 4 (Raiffa's figure 8.6 p 207). Three possibilities are depicted: (a) when one lottery is clearly superior to the other, (b) where the choice between the lotteries is not clear, and (c) where neither lottery is acceptable by itself, but by randomising across them it is possible to generate a jointly acceptable partition.

A.5 The Problem of the Panel of Experts

Suppose we are in a situation in which you and your panel of experts agree on the structure of the problem, but disagree on prior probabilities for uncertain states and on utility assignments for consequences. These respective assessments are as follows (Raiffa's Table 8.13 p 230):

EXPERT I					EXPERT II				
State	Act		Prior Probabilities		State	Act		Prior Probabilities	
	a_1	a_2				a_1	a_2		
Q_1	1	0	.2		Q_1	.5	1	.8	
Q_2	.5	1	.8		Q_2	1	0	.2	
Expected Utility	.6	.8	(1.0)		Expected Utility	.6	.8	(1.0)	

Thus, both experts (for different reasons) will recommend act a_2 as preferred to a_1. But if you, the decision maker, decide to decompose the problem into utilities and probabilities and split the difference between them on each element separately, you will come to the opposite conclusion, viz that a_1 is to the preferred to a_2. Thus:

State	Act		Your prior probabilities
	a_1	a_2	
Q_1	.75	.5	.5
Q_2	.75	.5	.5
Expected Utility	.75	.7	(1.0)

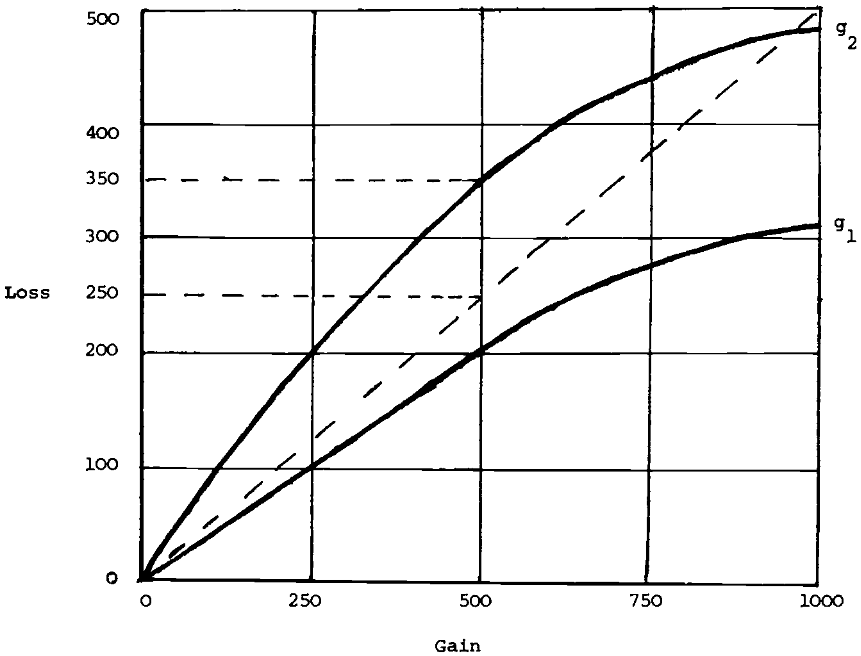


Figure 1

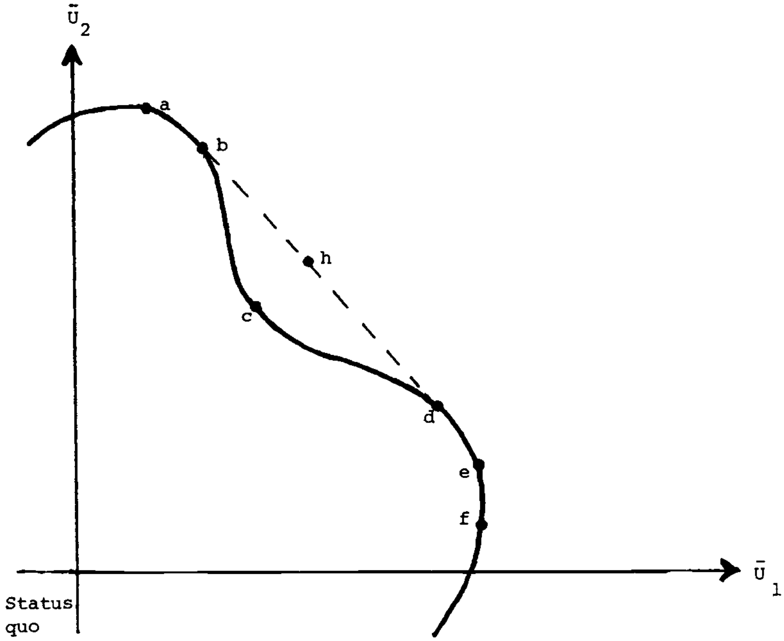


Figure 2

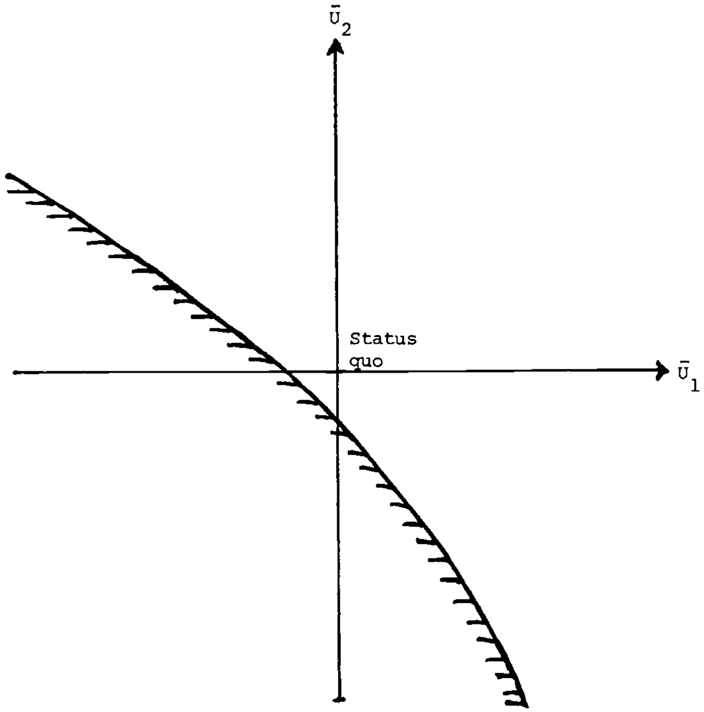


Figure 3

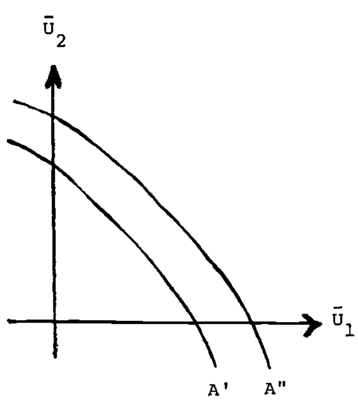


Figure 4a

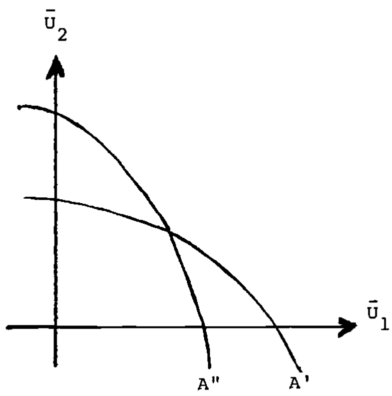


Figure 4b

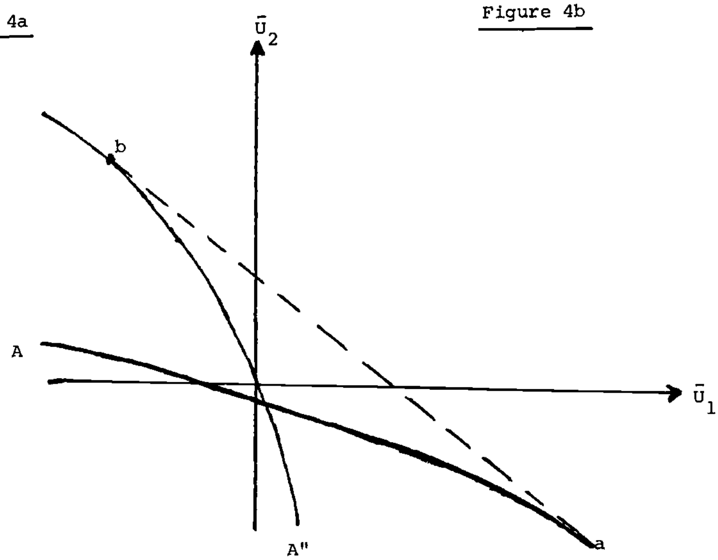


Figure 4c

GREGORY BAECHER
JACQUES GROS
KAREN McCUSKER

METHODOLOGIES FOR FACILITY SITING DECISIONS

Evaluating alternative sites for major constructed facilities requires comparing impacts of different levels and different types of established desirable yet feasible balances. Currently employed and proposed methodologies for evaluating the desirability of sets of impacts generated by large facilities are compared, and the theoretical assumptions implicit in each are discussed. In aggregate, the three sets of methodologies considered are Cost-Benefit Analysis and its various modifications, matrix or tableau methods of several sorts, and preference theory (of which utility is a special case). Primary attention if given to the structure of objective functions defined over impacts.

Authors' Preface

The question of siting decisions for major facilities involves complex interrelationships of spatial and societal distributions of impacts and at the upper end gradates into larger decisions of social policy and public welfare. We have attempted to isolate one facet of this process, the methodological approach to site evaluation, and analyse the assumptions implicit in commonly used or recommended methodologies. Were we considering an individual siting decision, we would attempt to use a combination of the techniques reviewed here as each has recommending properties that the others lack. Nevertheless, a discussion of each methodology by itself is helpful as it illuminates characteristics that might remain hidden in normal application. We have emphasised two seemingly simple concepts, which nevertheless are often transgressed in practice: rigorous properties of scaling, and interdependencies in desirability.

Our hope in formulating these thoughts stems from a desire not so much to advance the state of theoretical evaluation methodologies, as to aggregate a body of work in a consistent way so that site evaluation might be done without flagrant disregard for internal consistency and the principles of measurement.

As with any joint work, the responsibility and blame for the content of our observations are not equally shared. The organisation and writing of this review was primarily the work of G.B. Baecher; J.G. Gros contributed his ideas and experience with mathematical aspects of evaluation techniques and siting in general, and wrote some of the sections, and K.A. McCusker organised much of the literature, particularly that on matrix techniques.

We would particularly like to acknowledge the care which Harry Swain has taken in reviewing this paper and offering comments.

3. Methodologies for Facility Siting Decisions
by Gregory B. Baecher*, Jacques G. Gros & Karen McCusker

1 *Introduction*

Major constructed facilities generate a spectrum of impacts in addition to their central function: power plants generate air and water pollution, transportation projects generate land-use changes, and large water resources projects generate ecological disruptions. These impacts have always been recognised, if not before construction, then certainly afterwards. Historically, however, the central function of the facility has always received paramount attention, whether out of commitment to general welfare (the Roman aqueducts) or to profit (the Suez Canal). Secondary effects were considered of sufficiently lesser importance to be ignorable.

Large-scale water resources development during the first half of the twentieth century spawned increased attention to techniques of evaluating the spectrum of impacts generated by large facilities, but it has been the more recent difficulty of siting nuclear power facilities which has brought this problem to the awareness of the public. Often this awareness has manifested itself in emotional and at times semi-rational argument and confrontation. It would be unfair to attribute this widespread concern to greater vision and more complex times. Rather, our present attention stems from the growing scarcity of resources, in particular suitable sites for large facilities, and a growing affluence that allow us to adopt more multi-attributed definitions of societal well-being.

Ultimately, siting decisions are political, both in principle and in fact. Within the democratic framework they have traditionally been settled by debate, compromise and majority approval, constrained by notions of minority rights and long-term policy. However, the process of filtering large numbers of possible sites and making predictions about impacts is too large and burdensome for complete analyses in the political realm. This is where the analyst enters the siting decision process, and where the present review begins.

Analytical comparison of prospective sites requires balancing adverse and beneficial impacts against the multiple and often incompatible objectives of society. The coordinating theme of this balancing is the "desirability" we as a society associate with specific impacts against objectives, and this is what allows us to compare qualitatively different impacts of large facilities. Because it is the desirability of impacts and not their level that is important, decisions are ultimately based on subjective preference and not on "objective" criteria. One may elect, on subjective bases, to *use* a seemingly objective selection criterion -- for example, monetary cost -- but this does not make the selection objective; it rests upon the criterion, and the criterion upon judgement.

In approaching site selection, the analyst attempts to implement some consistent scheme for assigning desirabilities to individual impacts and for coalescing these into a decision. The result is a set of predicted impacts for each tentative site and each important objective, and two or three sites emerge which seem most favorable in the sense that the net desirability of associated impacts is the greatest. This short list of sites and the associated impact predictions (not the assigned desirabilities but the impacts themselves) is the departure point for political decision-making.

The nature of the results the analyst derives depends on the models (conceptual or mathematical) he uses to make impact predictions and the "consistent scheme" for eval-

---

* The senior author would like to acknowledge the support of the Rockefeller Foundation through its Conflict in International Relations Program Fellowship, RF 74025 allocation 21, during the tenure of which the present report was written.

uating and coalescing them. In this paper we compare these schemes in terms of the assumptions implicit in their structure and their applicability. We emphasise two points in this comparison:

1) Methodologies for comparing the desirabilities of impacts differ only in the specification of the objective function; this objective function makes implicit assumptions about the structure of desirability over impacts.

2) For scales of evaluation to be meaningful, one must know how numbers behave when combined by simple rules; any scaling and combination of impacts and their associated desirabilties must be firmly grounded in the theory of measurement.

Although this paper deals entirely with methodologies for evaluations, one should keep in mind that analytical evaluation is only one phase of the broad process of decision-making. By giving it preeminence there, we do not imply its actual preeminence in the entire siting process.

We carefully have drawn boundaries for our discussion so that primary attention could be focused upon methodological questions rather than political and social ones. One could easily argue that what has been eliminated is more important than what has been kept; we agree in spirit, but as always the normal constraints of time, expertise, and interest have dictated these boundaries. We assume that larger-scale policy decisions -- for example, whether or not a facility is to be constructed at all -- have already been taken; or alternatively, that larger-scale benefits and costs that are site-independent may be disassociated from siting itself. That is, the question whether a nuclear power plant or a highway should be built at all, while important and and issue of evaluation itself, is not considered here.

The paper is organised in four parts: siting decisions are discussed in general, then an overview of analytical evaluation schemes is presented along with their basis in measurement theory; three sets of methodologies are summarised and compared (cost-benefit analysis, matrix methods, and preference theory methods); and finally, application of the methodologies and general conclusions are discussed.

II *Siting and Public Decision-Making*

1 *The Siting Process*

On a conceptual level the question of siting is straightforward: it is merely the comparison of favourable and unfavourable impacts of a facility according to consistent rules for evaluating desirability, and the selection of the site that is found to have the highest net desirability. In reality, of course, this process is complex, involving both the seemingly irreconcilable interests of coalitions and vague notions of what social policy principles ought to be used as measuring rods of desirability.

The initial criterion in reviewing sites is feasibility. For a site to be feasible, the predicted impacts of placing a facility there must be within bounds chosen *a priori*. These constraints may include: excessive cost, excessive environmental degradation, undesired land-use alterations, and inequity in the distribution of net benefits. This process of eliminating infeasible sites is sometimes referred to as *screening*. Sites which remain after screening are evaluated in depth (Figure 1).

In the evaluation stage careful predictions are made of the type and magnitude of impacts generated by placing the facility at each feasible site. Desirabilities of individual impacts are evaluated as a function of the importance of the social objective they bear on, their magnitude, and their probability of occurence. This procedure rests on identifying social objectives and specifying desirabilities of impacts

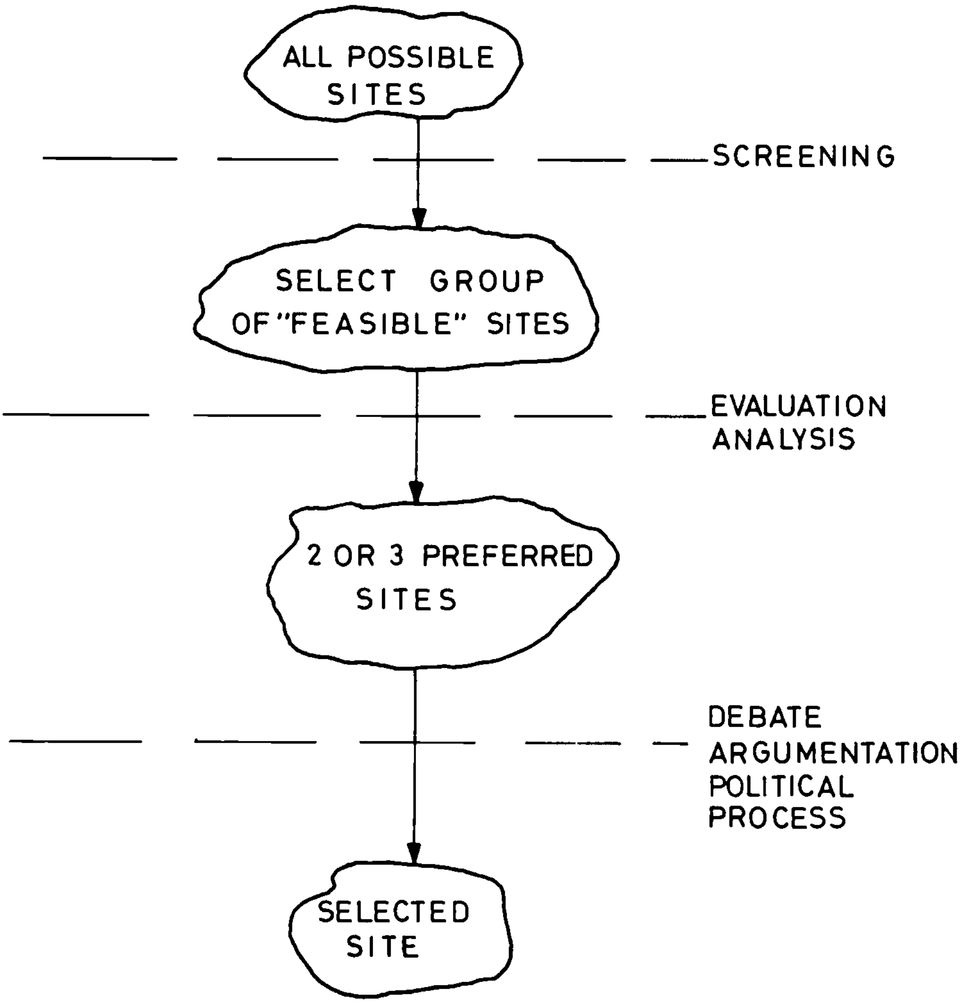

ALL POSSIBLE SITES
SCREENING
SELECT GROUP OF "FEASIBLE" SITES
EVALUATION ANALYSIS
2 OR 3 PREFERRED SITES
DEBATE ARGUMENTATION POLITICAL PROCESS
SELECTED SITE

FIGURE 1

against those objectives. Impact predictions, while often difficult to make with precision (Buehring, 1975), present technical rather than philosophical problems; whereas the central questions in evaluation, and those on which the entire analysis depends, are what social objectives are used for evaluation, and whose concept of desirability is adopted?

2 *Social Welfare and Selectivity*

Ideally, one would like to make decisions having a social impact in light of a general theory of social welfare using a comprehensive objective index, which is based on the ethical or normative precepts of the society. In reality, of course, attempts to develop a social welfare function have not been fruitful, so in practical decision-making a more pragmatic and less "objective" criterion must be reverted to.

Our ability to make comprehensive evaluations is limited not only by lack of a general welfare function, but also by our inability to predict the myriad of secondary, tertiary, and higher-order impacts which a decision generates. The 1960's thus saw the development of large simulation models, many of them for regional planning, whose purpose was to simulate interactions and dependencies via a logical chain too complicated to be analysed intuitively. The hope was that this approach to analysis would enable us to predict indirect impacts and include them in decision-making. But this attempt, too, has not been entirely successful (Lee, 1973; Brewer, 1973).

This brings into clear perspective the problem of identifying and selecting important impacts for analysis. We must select a limited number of objectives against which we consider impacts to be important, and a limited set of indices for prediction. In assigning desirabilities to levels of those impacts, we must do so subjectively -- if not in the way the final numbers are placed in impact levels, then in the way assumptions are made and data collected. The criteria and measurement represent value judgements by the analyst whether or not he readily admits it. There is a continuum between the analyst and the political decision-maker. In both cases decisions are made the same way: the analyst tends to use a larger criteria set, and explicitly combines his evaluations according to logical rules. But the philosophy of decision and the form of evaluation are the same at their philosophical foundations.

What is the overall criterion of evaluation? Given benevolence in government or a democratic ethic, the criterion of evaluation is the well-being of the population. In positive economics and democratic theory this is held to be the preferences of individuals within society. How these preferences are assessed and interpreted is integrally related to the technique used for comparing desirabilities of impacts against objectives. Assessment methods may be indirect as in using market structure and prices, or direct as in opinion surveys. Once again, the analyst's role in this process is to interpret those preferences from data and logically combine them so as to yield recommendations for the political decision-makers (who ultimately interpret desirabilities judgementally).

3 *Siting vs. Planning Decisions*

National and regional planning goes on at many levels, and it is in the analyst's interest not to confuse the proper distribution of authority and decision responsibility within that hierarchy. Not every decision made in society must involve a reassessment of the basic ethical and economic policies of society. In other words, the decision to site a nuclear power plant is not the most appropriate point for reassessing national energy policy. In actuality, the siting decision may be the only (or most accessible) point at which a citizen may exert pressure against what is perceived as an unresponsive political process; but from the point of view of governmental planning this is clearly not the case. On the other hand, though, gradations of planning responsibility are fuzzy, and the resources for analyzing major siting decisions may be much greater than those available for planning overall regional development; perhaps this is an inverted situation, but it is nevertheless the case. So, another facet of the selection question is, how broad does one make the impacts and societal objectives considered, and where in the analysis does one adopt the results of higher-level decisions as either constraints or scales of desirability? In a hierarchy of decision which is not rigid, this question assumes considerable importance.

Ostensibly, we have planning authorities whose business it is to evaluate proposals for regional development and to arrive at preferred scenarios. To the degree that such bodies do have sufficient expertise and financial resources to accomplish their mandate, impacts generated by siting a facility should be evaluated for their compatibility with these preferred plans. If the preferred plan calls for slow development and primarily agricultural patterns of land use, then a facility causing inharmonious land uses (e.g. large transportation facilities) generates undesired development impacts. In the reverse situation, a facility inducing larger local employment, and thus accelerated development, would be deemed more desirable than one that does not. In this ideal world the siting analyst's life would be simpler.

When no local planning authority exists, the ethical question arises, is it appropriate that the analyst treat questions of regional development policy. If such questions have not been dealt with, they *de facto* become his responsibility, and he must grapple with them. Conceptually, the task is clear, but practically it is difficult; the project's long term indirect impacts on population, migration, settlement, and regional land use must be considered in the same way as are impacts against other objectives. Typically, this can only be accomplished by judgemental or conceptual models, or by rather large computer models which include complex interactions of employment, infrastructure development, and changes in environmental quality. The latter models suffer the disadvantages of all large models as discussed by Lee (1973).

If longer-term predictions of land-use and development impacts can be made, the analyst is still faced with the problem of evaluating the desirability of such changes. The time is past when simple economic indices of regional development (e.g. increases in tax base, increases in real income flow) can be used as positively correlated measures of desirability. At present even the desirability of regional development is in question. Local residents do not always favor increased development; or they may do so, while far distant urbanites prefer to maintain unspoiled rural landscapes -- even if they are unlikely ever to visit the region.

In short, unless a well-covered plan for regional development exists, the analyst, by default, must develop a surrogate plan. We would hold that this is not really his mandate, but a burden which is dealt him.

4 *Coalition and Equity*

There are two distinct concepts with respect to the disaggregation of society into groups. The first is that individuals place different weights on the desirabilities of impacts and on marginal rates of preferential substitution among impacts; here, the question naturally arises whose definition of desirability ought to be used in siting decisions. The second concept is that of the distribution of benefits and costs over society. Large facilities have uneven spatial and social distributions of impacts, and one may value a level of equity in these distributions. We will address legitimacy of interest first, and then return to equity.

Welfare economics has attempted to structure a theory to account for differences in individual preference, and has succeeded mostly in proving the great difficulty or the impossibility of doing so. Pragmatically, therefore, in siting decisions one normally views differences in preference or definition of desirability as being represented by groups of opinion. While the term is misused in this context, we often call these assumedly homogenous clusters of preferences interest groups, and we assume that the interests of individuals within groups can be approximated by a single structure of desirability for impacts. (In fact, this is not the case; interest groups either are not organised groups at all (e.g., see Olson, 1965), or are coalitions formed for attaining some common goal, but one sought by each individual within the coalition for perhaps very different reasons.) Such simplifications are undertaken

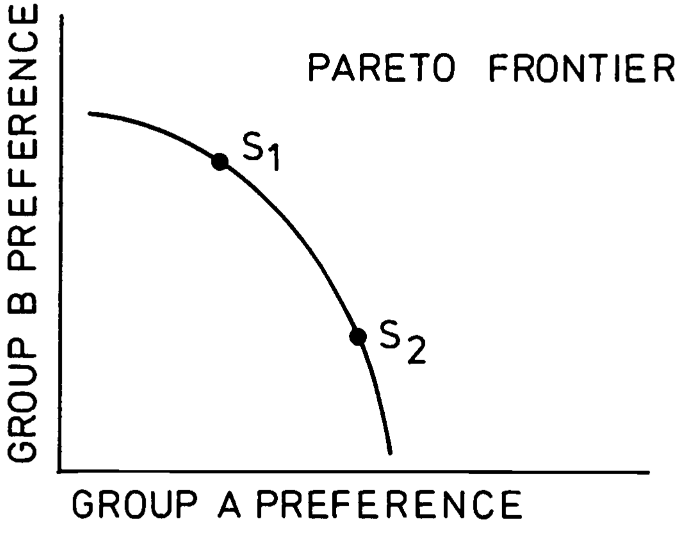

PARETO FRONTIER
S1
S2
GROUP B PREFERENCE
GROUP A PREFERENCE

FIGURE 2

to make the problem of analysis tractable, just as one makes simplifications in analyses, whether they be mathematical or purely judgemental constructs.

The ultimate burden in combining different concepts of desirability rests with political decision-makers, this being a fundamental function of the political system. The analyst's role is to indicate to the political decision-maker the implications of weighting different groups' interests in different ways on the "optimal" decision. In the more purely economic approaches to siting decisions, such as cost-benefit analysis, an assumption is made that differing preferences are naturally and properly aggregated in the market-place; yet even here, the desirablities of non-market impacts (or impacts with which there is little experience) still require an artifical weighting and coalescence. If one uses the market-aggregated willingness-to-pay of urban and rural residents as a measure of the desirability of aesthetically pleasing landscapes, a value assumption is still made about the relative weights given each group, through the weights are not explicitly stated as they would be with other methods. No matter how a siting decision is evaluated, the preferences of different groups must be weighted. Methods that do not do so explicitly must do so implicitly; usually this means weighting all groups equally.

Conceptually, one can think of the question of weighting interest group preferences as movement along the so-called Pareto frontier. This surface is the locus of all decision alternatives (sites) for which no other alternative exists that would be equally preferred by all groups and more preferred by at least one. In Figure 2 no sites are available which, for the several impacts they generate, are more preferable to *both* groups A and B than, say, site # 1. Here, we would hold that it is the analyst's role to determine those sites which are on the frontier, and the sensitivity of each group's level of desirability to movement along the frontier. The decision among sites on the frontier is innately political, although this task might be aided by sensitivity analysis which would indicate "optimal" sites for ranges of weights applied to each group's interest.[1]

The dynamics of the political process makes the view just presented myopic. At any one time many projects are being considered by political bodies, and often equity is achieved not within a single project but over several projects. A project that favours one interest group over another might be offset by one which favours in reverse. In the democratic framework this is related to keeping constituencies satisfied (or placated) and is a natural offshoot of the legislators' self-interest in remaining in office. Thus, the question of whose measures of desirability we use is closely related to the concept of equity of impact distribution.

A fundamental tenet of contemporary political philosophy is that fruits and labours of society should be equitably shared by members of society. However, equity is one of those nebulous policy concepts mentioned in the introduction. No one is quite sure, in operational terms, what equity ought to mean, but we all know that it's important. In traditional project decision-making, equity has been treated either as a prior constraint that a proposal must satisfy or as an "external" weighted in conjunction with economic efficienty. A project that is otherwise efficient in the sense of producing a net increase in benefit to society, irrespective of to whom it accrues (i.e. potential Pareto improvement), might be discarded if it produces what is politically viewed as a severely adverse distribution to those costs and benefits. More recently introduced methodologies, as discussed in Sections IV - VII, attempt to measure equity explicitly as one impact of the decision and subjectively assign desirabilities to it which can then be combined with other impacts. We are, however, far from a workable definition of equity or attribute scale that could be included in an analysis; even

---

1. The concept of Pareto optimality and the frontier are used here for illustration only. There are theoretical questions relevant to using Paretian analysis in actual decisions, one of which is taken up in the appropriate Appendix.

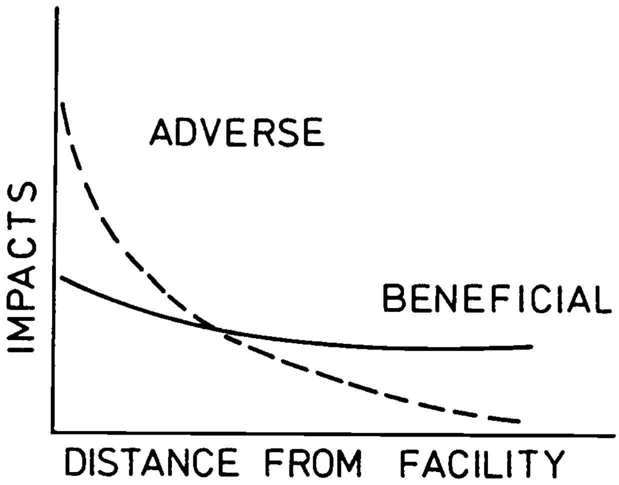


FIGURE 3

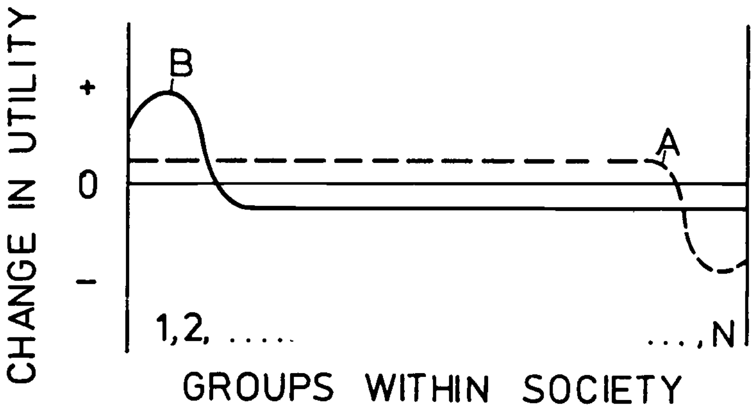


FIGURE 4

equity of income distribution generated by projects, a seemingly simple problem, is difficult to grapple with normatively (Mishan, 1971). The further complication, in siting studies, of the geographic distribution of effects (Figure 3) makes the problem exceedingly difficult unless purely judgemental approaches using political opinion are introduced.

Once again, though, to maintain our perspective merely at the single project level is naive. Political decision-makers almost invariably favour projects generating impact distributions as shown schematically by curve A in Figure 4 over those generating curve B, even though an analytic index of equity might rate A and B at about the same quantitative level of "inequity."[1] There is a quality difference in the inequity caused by A and B because, if need be, a purely redistributional project can be formulated, aimed directly at the groups adversely affected by project A. In a conflict resolution sense, this would be the same as a side payment to adversely affected groups to "get them to go along" with the project -- something which is not at all rare in siting major facilities. Techniques used to include equity in specific evaluation methodologies are discussed further in Sections IV - VI.

5 *Temporal Distribution of Impacts and Irreversibilities*

Benefits and costs accrue from a project non-uniformly in time. Capital outlays for facility construction are necessarily made at the very beginning, while financial returns on investment, social disruptions, and environmental impacts come at varying times, from almost immediately to the distant future. Some irreversible impacts, such as major ecological changes, continue in perpetuity. Ideally one would like some analytical way of treating these streams of benefits and costs.

Analytically, this evaluation might be simply represented by a series of the following type, in which NB_t is the net benefit of the project accruing at time t:

$$NB_{total} = NB_0 + v_1 NB_1 + \dots + v_n NB_n . \qquad (1)$$

The question is how to evaluate the constants $v_1, \dots, v_n$; and whether or not the aggregation ought to have a more complicated form than a simple sum (Meyer, 1969; Koopmans, 1960). This is a problem that has received extensive attention, yet remains unanswered.

The traditional way of handling intertemporal streams of costs and benefits has been to assume an additive form as shown in Equation 1 and adopt a discounting factor relating the value v_t to its predecessor by a constant ratio, r,

$$\frac{v_{t-1} - v_t}{v_t} = r = \text{discount rate.} \qquad (2)$$

Koopmans (1960) gives the necessary conditions for this form of discounting, called the "discounted sum", to be theoretically correct. The discounted sum has been generally applied in cost-benefit analysis, and considerable work has gone into techniques of establishing appropriate discount factors (e.g., Layard, 1972; Roskill, 1970; Mishan, 1971, UNIDO, 1972). Some of these are the market interest rate on capital, the marginal rate of productivity of capital in the economy, or simply a value judgement of political decision-makers. The time aggregated net benefit (NB_{total}) of a project may fluctuate substantially on the basis of changes in the discount rate, and varying of this rate has often been used to justify bureaucratically favoured projects that would not be justified by more impartial analysis (Berkman and Viscusi, 1973). Further, the normal procedures for establishing the discount rate are not entirely satisfactory because for societal projects, the discount

---

1. This example is due to H. Swain (personal comminication, 1975).

rate reflects social policy on how much one is willing to forego now for future benefit. In a traditional sense, the best procedure, as with equity, is to do a sensitivity analysis using discount rate as a variable, and then see how high or low the rate would need to be to change the "best" decision.

Specifically with respect to siting decisions for large facilities, two points are important. First, many of these decisions are private ones involving private funding; this being the case, the discount rate for financial costs and returns can be chosen by the private agent and will probably reflect market costs of capital. Second, the siting decision as we have outlined it here is not a decision to construct or not to construct a facility, but is limited rather to where to construct it. Therefore, as a given type of facility constructed in different places generates approximately the same temporal distribution of impacts (although not in the same intensity), siting decisions are less sensitive to discounting than the overall project decision might be.

While discounted sum techniques may be appropriate for financial impacts even though the actual rate of discount is difficult to specify, the discounted sum is not so apparently appropriate for non-financial impacts (i.e., social and environmental ones), and the whole question of non-renewable resources is still in an embryonic state of analysis. An approach of the type used by Meyer (1969) may shed light on time streams of non-financial impacts as that work expands; similar comments can be made on work evaluating alternatives that exhaust non-renewable resources or generate irreversible impacts that is being undertaktn by Krutilla and his associates at Resources for the Future (Fisher and Krutilla, 1974; Krutilla et al., 1972). At present, these remain unanswered questions.

An associated set of problems is that of option fore-closure, resilience, and incrementalism. One type of irreversibility, although not the type usually dealt with, is that of foreclosing options that might later have been open. Krutilla et al. (1972) discusses this, as does Walters (1975). Option foreclosure means that impacts generated by a decision will make future decision alternatives impossible. For example, siting a nuclear waste storage facility will mean that the site is forever unusable for other purposes. The degree of desirability of foreclosing future options depends on the probability that one would at some later time elect to use them, the time when that might occur, and the benefit that would have been derived from their use. In some cases, positive discounting factors (i.e. which give more weight to future benefits) might be appropriate to describe goods that will become increasingly scarce with time. Some of these might be open space, environmentally undisturbed wilderness, or non-renewable resources (Krutilla, 1972). Option foreclosure also deals with impacts that cannot be predicted, but that will change the environment of future decisions and thus change in unpredictable ways the options that would have become available (Walters, 1975). Perhaps the best way of treating such foreclosure practically is by instituting incremental decisions the results of which can be sequentially evaluated, and by designing alternatives which are resilient to unforeseen events. In siting, while incrementalism can be practiced only by building small facilities, resilience would mean selecting sites that are far enough removed from population, naturally undisturbed areas, etc., that unforeseen impacts would have little undesirability. Unfortnately, it is because of a lack of such sites that the issues has become so important.

A major issue growing out of resilience and option foreclosure is what Hafele has called "hypotheticality", that is, the problem of dealing with low-probability events with which we have no experience, (e.g., large-scale accidental releases of radiation from reactors) (Hafele, 1974). This problem increases in importance with rapid technological developments which exclude an incremental approach to decision-making. The question is not beyond the bounds of the siting decision since the major objection to urban sites is large-scale health and safety risks.

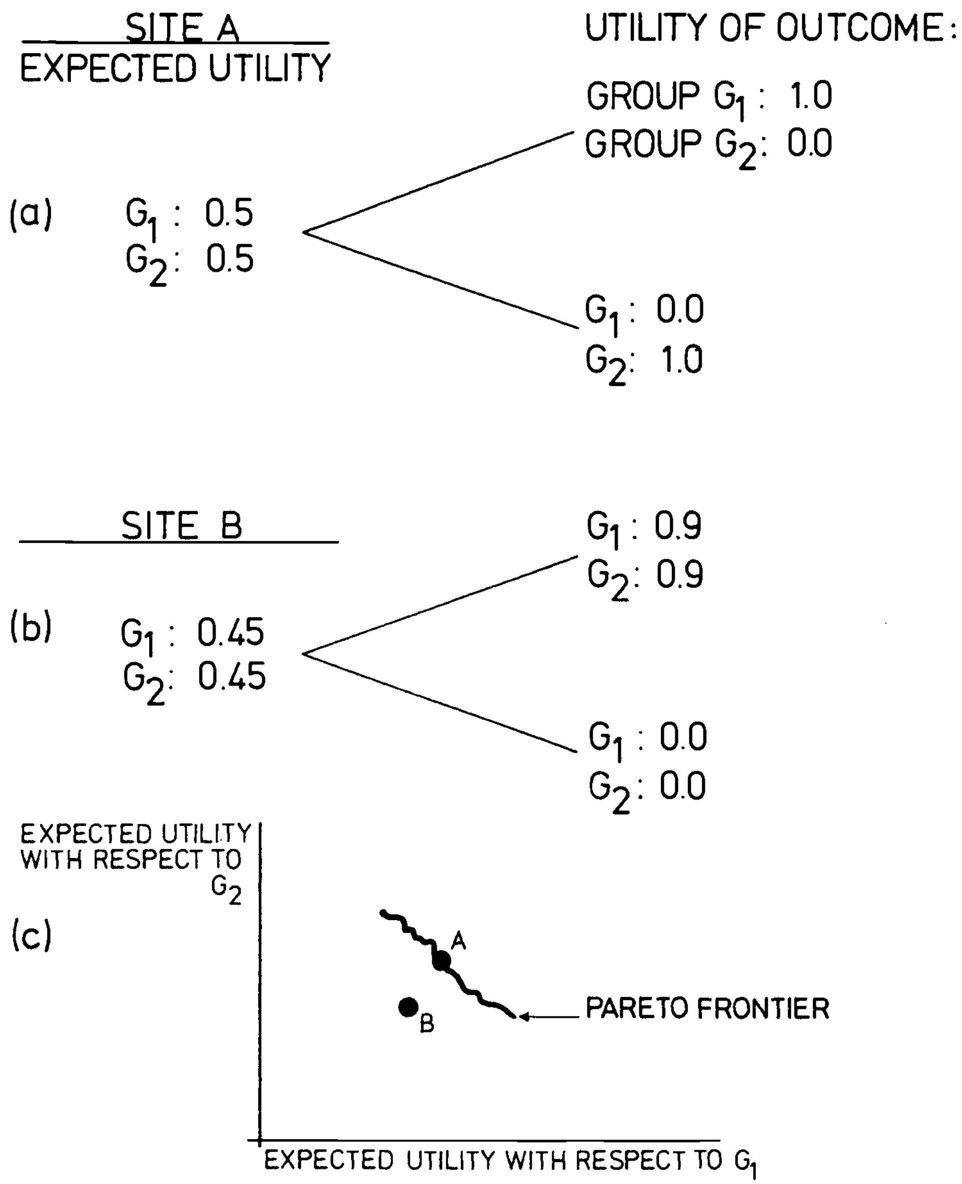

SITE A
EXPECTED UTILITY
UTILITY OF OUTCOME:
GROUP G1 : 1.0
GROUP G2: 0.0
(a)
G1 : 0.5
G2: 0.5
G1: 0.0
G2: 1.0
SITE B
G1 : 0.9
G2: 0.9
(b)
G1 : 0.45
G2: 0.45
G1 : 0.0
G2: 0.0
EXPECTED UTILITY WITH RESPECT TO G2
(c)
A
B
PARETO FRONTIER
EXPECTED UTILITY WITH RESPECT TO G1

FIGURE 5

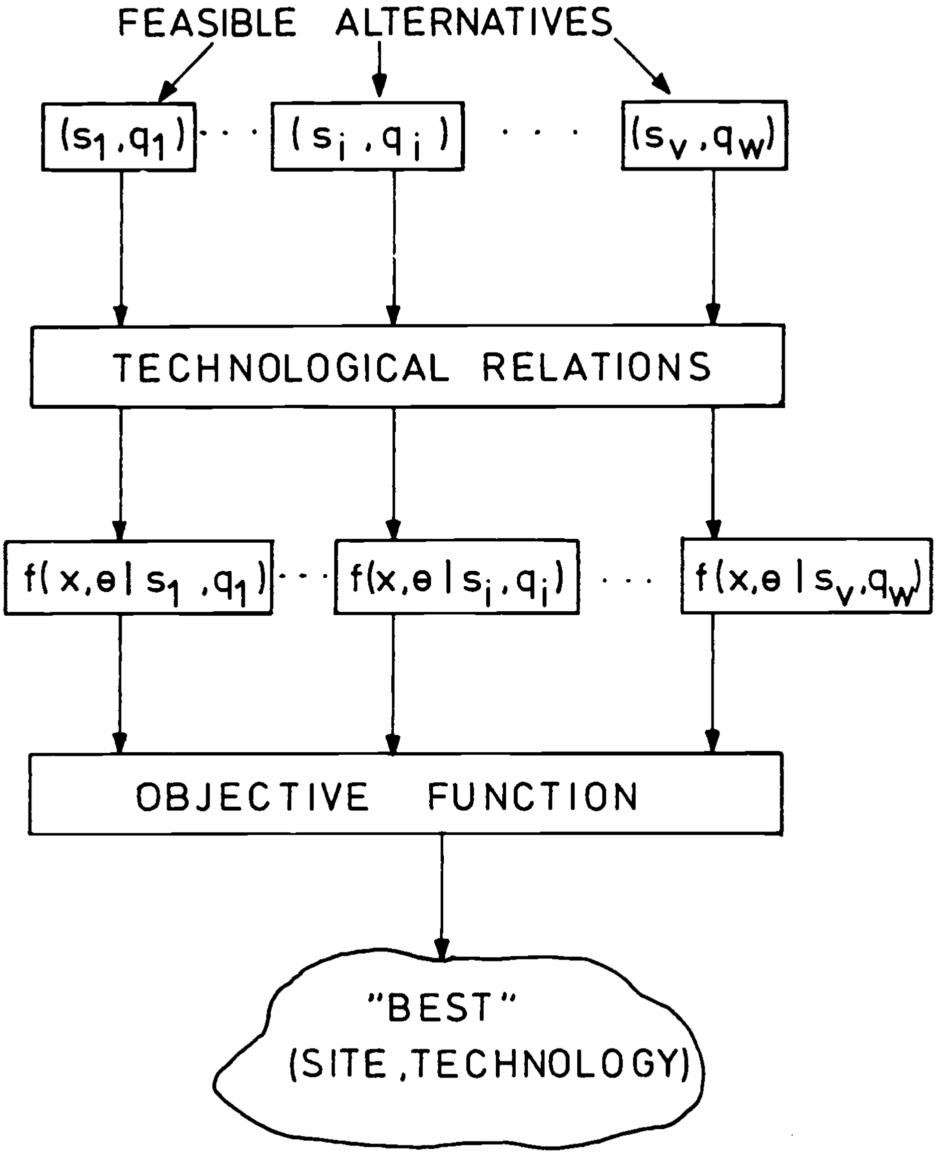

FEASIBLE ALTERNATIVES
(s1 ,q1) · · · (si ,qi) · · · (sv ,qw)
TECHNOLOGICAL RELATIONS
f(x,θ | s1 ,q1) · · · f(x,θ | si ,qi) · · · f(x,θ | sv,qw)
OBJECTIVE FUNCTION
"BEST"
(SITE ,TECHNOLOGY)

FIGURE 6

APPENDIX

Pareto Admissibility under Uncertainty

If equity is considered important by the decision-maker, an optimal alternative need not lie on the Pareto frontier defined by interest-group preference (Keeney, personal communication). In the case shown in Figure 5 the problem is to select site A or B. These sites are associated with uncertain impacts along one attribute which lead to different levels of desirability (i.e., utility) for the two groups G_1 and G_2. Clearly, alternative A is a point on the Pareto frontier composed of the expected utilities of impacts, and has a higher expected utility for both G_1 and G_2 than alternative B, which must therefore be below the frontier. Yet, if the decision-maker considers equity to be an important attribute of any set of impacts, then he might favour alternative B to A, because no matter how impacts accrue, equity of impact will be maintained. Thus, under uncertainty an optimal decision alternative need not be on the Pareto frontier.

III *Structure of Evaluation Methodologies*

Analytically, all evaluation methodologies have a similar structure. In this section we discuss that structure and introduce terms and notations to simplify our further comments.

Siting decisions are, in fact, decisions among variables in two sets: a set of possible sites, and a set of possible facility technologies. Jointly, these might be called the set of *feasible alternatives*. Symbolically, if the set of sites is $\underline{S} = [s_1,\ldots,s_v]$ and the set of facility technologies is $Q = [q_1,\ldots,q_w]$ then the set of feasible alternatives is composed of all possible pairs (s_i,q_j) that remain after screening. As impacts depend on both the site and techology selected, "siting" decisions must involve both variables.

Feasible alternatives are judged by their impacts against a set of objectives society holds important -- e.g., cost, environmental degradations, and social disruption. Since objectives are usually vague and qualitative concepts, a set of indices is chosen for measuring levels of impact against objectives. We will call these *attributes*. For example, to quantify the degree of impact a site-technology pair has on the objective "minimize water pollution," we might use the attribute "concentration of pollutant y in effluent waters." Associated with each objective is at least one scalar or vector attribute; let the set of attributes be denoted $\underline{X} = [x_1,\ldots x_n]$.

Decisions are made on the basis of predicted impacts measured on the set of attributes associated with important objectives. These predictions are made judgementally by experts using mathematical and statistical models, basic concepts and relationships from the physical and social sciences, and the like. In general, these predictive relationships might be said to map site-technology pairs onto the attribute space. Since predictions are uncertain and depend on exogenous random variables, such as weather, accidents, and future population densities, they are actually probability distributions defined over the set of attributes. Collectively, we call these predictive distributions the set of *technological relations*, and denote them as the joint probability function

$$f(\underline{x},\underline{\Theta}|s_i,q_j) \quad ,$$

in which $\underline{\Theta}$ is the set of exogenous variables (Figure 6).

Implicit in the set of technological relations are not only impact predictions for a given site-technology pair, but also the *marginal rate of technical substitution* among impacts; that is, the rate at which it is technically possible to trade one impact for another (in an uncertain domain). For example, pollution emissions can be reduced if one is willing to increase project cost; or a natural wilderness area can be preserved if one is willing to site a power plant nearer to a densely populated area. The concept of marginal rate of technical substitution is an important one because it is, in

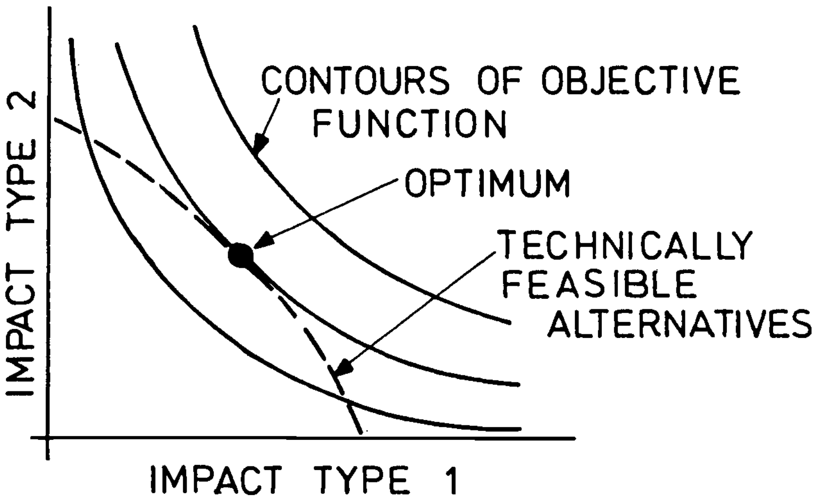

CONTOURS OF OBJECTIVE FUNCTION
OPTIMUM
TECHNICALLY FEASIBLE ALTERNATIVES
IMPACT TYPE 2
IMPACT TYPE 1

FIGURE 7

some sense, half of the evaluation. The other half is the *marginal rate of preferential substitution,* the rate at which one impact can be traded for another without changing the aggregate level of desirability of the set of impacts. At the optimal decision (under certainty) these two marginal rates are equal (Figure 7).

The marginal rate of preferential substitution is implicit in whatever objective function is used to evaluate different sets of impacts. Objective functions are numerical representations of preferences for different attribute levels; the optimal decision is the one which has the largest objective function value. It is the nature of this objective function and of the assumptions implicit in its derivation which distinguishes evaluation methodologies from one another, and which is the focus of the present review.

1 *Objectives*

It is assumed here that objectives for siting decisions are known or can be generated. Some of these objectives are "to provide adequate service", "to minimise environmental degradation", "to minimise social disruption", and "to minimise adverse health and safety effects". Most of them can be identified on the basis of past decision-making (or the criticism of that decision-making) and from the siting literature. Certainly an extensive list of impacts that might be (and for nuclear power plants in the United States, must be) accounted for appears in *USAEC Guide 4.2* (1973).

The set of objectives should have several properties: it should be *complete,* in the sense that it contains all important considerations on which a decision has impacts; it should be *non-redundant* in the sense that "double-counting" is minimised, and it should be of *minimum size* to facilitate analysis.

Hierarchies of objectives exist; it is only at the lowest level that objectives become specific enough for one to grapple with them analytically. At high levels are such objectives as those cited above, which are too abstract to use in an actual decision. In constructing their hierarchy, one attempts to structure objectives so that each highest-level objective comprises sub-objectives which fully describe its important aspects and yet can be dealt with more straightforwardly. For example, within or below the objective "minimise environmental degradation" might be the sub-objectives "minimise adverse impact on aqueous life forms", "minimise adverse impacts on terrestrial life forms", and "minimise aesthetic degradation of landscape and adverse aesthetics of water and air pollution" (Figure 8). Specification of sub-objectives not only facilitates analytical treatment, but also *clarifies* and *defines* the upper-level objective for the purpose of analysis. Thus care must be taken to assure that the substrata of the objectives hierarchy do actually meet the intentions of the analyst or decision-maker. One mechanism for constructing the objective hierarchy is to ask whether or not sub-objectives *do* completely describe upper-level objectives, and if they do not what additional sub-objectives must be provided so that they do.

It is not our purpose there to dwell on the question of how inclusive or finely divided the objectives hierarchy should be; this problem is treated elsewhere (e.g., Manheim and Hall, 1967; Keeney and Raiffa, 1976). Certainly, however, all sub-objectives that may change the result of analysis must be included, although sometimes they may be treated in sets to facilitate quantification (Ting, 1971). In the end, the point at which formalisation stops is a judgemental problem.

2 *Attributes*

Since objectives, even at lower levels in the hierarchy, are usually not measurable concepts, indices must be specified over which impacts can be scaled; these are called attributes in the present paper. Given the sub-objective "minimise thermal pollution to receiving waters", a typical attribute might be "increase in temperature of receiving waters in degress centigrade". Listing of typically applied attributes may be found in USAEC (1973) and in Keeney and Nair (1974). With each lowest-level objective some attribute is associated, which itself may be either a scalar or vector.

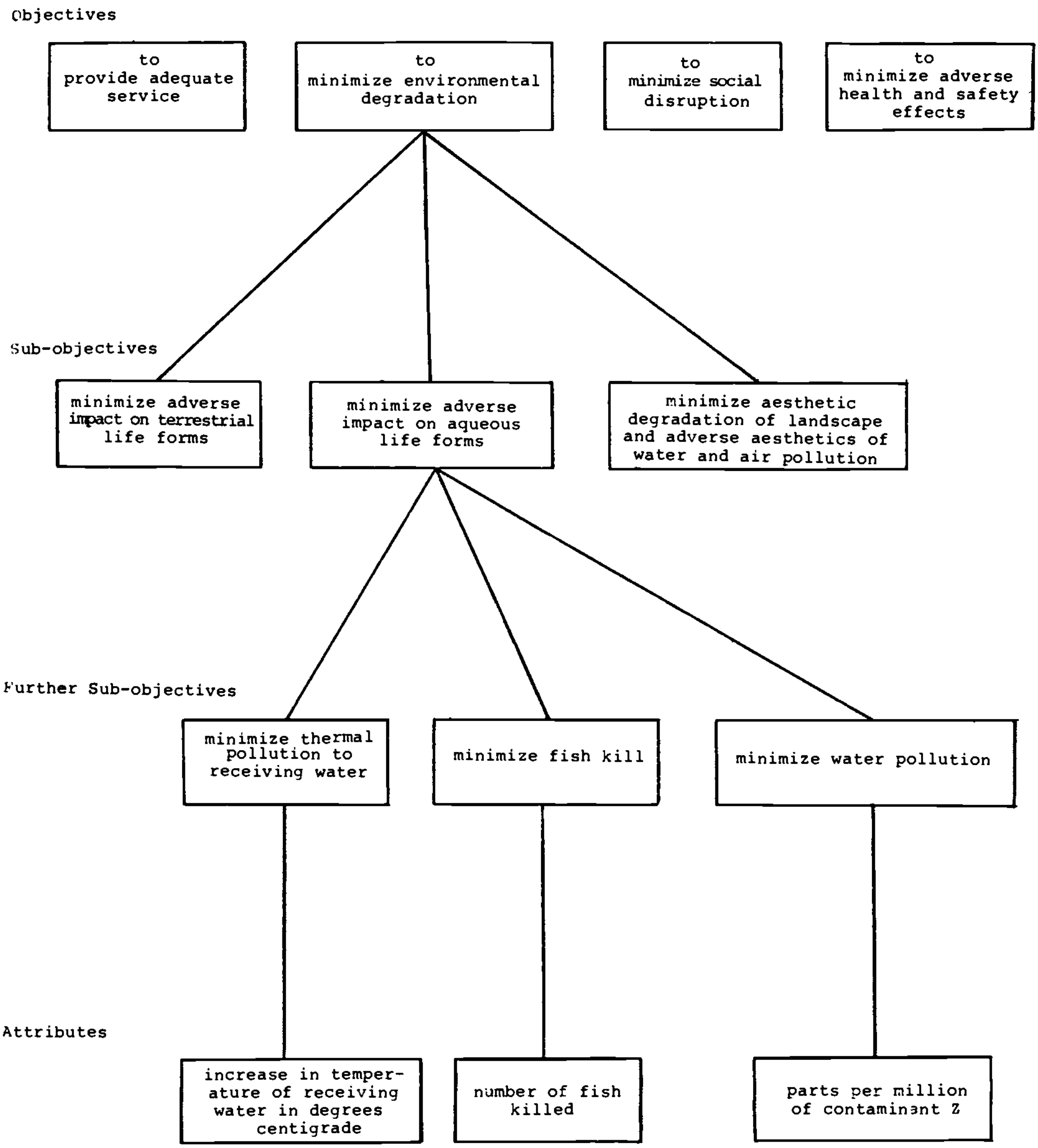


Figure 8

Example of an Objectives Hierarchy

Individual attributes must be, in the terms of Keeney and Raiffa (1976), *comprehensive* and *measurable*. Comprehensiveness is the property that the level of impact as measured on an attribute fully expresses the degree to which the associated objective is achieved; measurability is the property that predictions can be made about the impact of a proposed site and technology alternative in terms of that attribute, and that the objective function (i.e., desirability) over values of the attribute can also be assessed.

The *set* of attributes should also display two properties, *non-redundancy* and *minimum size*. The set should be non-redundant so that impacts are not double-counted (e.g., see McKean, 1958) and of minimum size for analytical tractability.

The set of attributes associated with the objectives hierarchy and each attribute itself do not uniquely follow from the objectives, and only with a small fraction of the objectives considered do attributes immediately suggest themselves. Thus the selection of attributes may itself affect the outcome of analysis; one is well advised to proceed with great care and to assess retrospectively the sensitivity of analytical results to attribute selection.

Attributes that do follow immediately from an objective are said to be *natural* attributes. For example, if one sub-objective were to "minimise 'fish kill'", a natural attribute would be "number of fish killed". When an attribute does not follow immediately from the objective, as is normally the case, a *proxy* or *surrogate* attribute must be employed. For example, one might associate the attribute "parts per million of chemical contaminant Z" or "BOD" with the objective "minimise water pollution". These are not direct measures of the water quality the associated objective deals with, but rather are correlates, and may be chosen either because they primary property is inherently unmeasurable or becuase the natural measure is analytically intractable. To specify water pollution adequately, for example, would require a vector attribute of large dimension, so large that it could not be used in analysis.

A second reason for choosing a proxy attribute is that data may be more easily obtainable for it than for an attribute that seems to follow more naturally. This may be due to ways in which data have been historically collected or aggregated, because certain types of monitoring are cheaper or quicker than others, or because it is easier to specify the objective function over some attributes than others. In cost-benefit analysis and other methods which use money as a measure of desirability, this increased ease may arise because some attributes have closer analogs in the market place than others; and in methods such as utility analysis which use subjective valuations of desirability, because individuals find it easier to think about certain measures of impacts than about others.

In siting problems impacts arise for which even proxy attributes cannot be identified, either because adequate indices have yet to be developed for very complex phenomena, or because the impact seems inherently non-quantifiable. In such cases scenarios are often specified in qualitative terms and values of desirability assessed directly over the scenarios. This technique is receiving increasing attention in problems of facility siting, particularly with aesthetic impacts such as visual quality of the landscape (Jones et al., 1974; Burnham et al., 1974). At present these approaches generally specify a rating scale associated with adverbal descriptions and scenarios, rate impacts of contending alternatives along that scale, and subsequently assign desirabilities to the scale. As this work proceeds, proxy attributes or scales may be developed which better lend themselves to quantified description (Holling, 1973).

Money is often taken as an attribute with which to measure the impacts of site technology pairs. Indeed, with such methods as cost-benefit analysis there is a strong bias towards expressing as many impacts as possible in monetary terms since impacts are coalesced in monetary units. There is nothing improper about this approach, as long as impacts can be readily and comprehensively expressed in monetary units. Often, however, money is used not as the attribute of impact, but rather as the measure of desirability of an impact which is itself measured along another scale --

Level of Scaling
Linearity of Desirability over Individual Impacts
Independence or Non-independence among Impact Desirabilities
Analytical Treatment of Uncertainty
Marginal Rates of Preferential Substitution
"Objectivity"
Explicit Aggregation

Figure 9

Characteristics of Objective Functions

for example, a monetary value is assigned to each fish killed by pollution. As desirability may be expressed in any consistent unit, again there is nothing innately improper in this approach. However, some units, such as money, may have interrelationships within the measure itself which are not shared by whatever one is trying to measure; the analyst must be careful that properties of the measure not reflected in the phenomenon are not employed in the mathematical analysis. This is an important point which will be developed later in this section.

3. *Objective Functions*

We have already said that the distinguishing characteristic of evaluation methodologies is the form of the objective function. We now turn attention to properties of objective functions that distinguish one from another. Figure 9 lists these properties.

Desirability of an impact may be measured to an ordinal, interval, or ratio scale, (a brief review of scaling theory is presented in Appendix III). Admissible operations on measurements of desirability depend on the scale used. If desirability is measured to an ordinal scale, as with some matrix methods, then the operations of addition and multiplication necessary for aggregation are not permissible. Thus, aggregating ordinal data yields numbers whose relationships to one another have no meaning. If desirability is measured to an interval scale, then ratios of desirable to adverse impacts have no meaning. One is generally reticent about making stronger assumptions than one must, but practical advantage can be realised by defining desirability to a higher scale than is theoretically necessary. Decisions among alternatives having multi-attribute but deterministic impacts require only that desirability be measured to an ordinal scale, and in fact Major (1974) has done so in water resources location problems. In practice, however, it may be much easier to assess and computationally handle desirability if it is measured to an interval or ratio scale. Of course, this ease of application is bought with more restrictive assumptions.

The level of scaling to which impacts are measured and that to which desirability is measured need not be the same. For example, financial costs of a project are measured in monetary units, that is by a ratio scale, yet the desirability of levels of cost may be only an interval measure. On the other hand, impacts such a visual aesthetics may be measured only to an ordinal or even nominal scale, yet the desirability may be measured to an interval scale, or even a ratio scale (e.g., "willingness-to-pay").

Given an interval or ratio scaling for desirability over one attribute, the objective function may be linear or non-linear (Figure 10). Assuming that each increment of impact is just as important as every other increment leads to linearity, as when one assigns a unit cost and multiplies by the number of units. Linearity means constant marginal rates of changes of desirability with unit increases in impact.

The desirability of impacts measured over multiple attributes may be either independent or non-independent. Stated another way, the level of desirability of an impact versus other impacts may or may not depend on the levels of other impacts. For example, the decrease in desirability caused by a unit increase in project cost may or may not depend on the level of environmental impacts. If the unit cost increase is considered less important for a project with very low environmental effects than for one with high environmental effects, then the desirabilities are non-independent; they do not follow the relationship

$$D(\text{cost, environmental effects}) = D(\text{cost}) + D(\text{environmenal effects})$$

Independence among the desirabilities of impacts must be distinguished from technical independence among them. Two impacts such as visual aesthetics and heat release may be technically independent in that the beauty (of lack thereof) of a facility might play no part in the level of pollutants released, or vice versa; while the marginal desirability of increases in pollution may depend on the desirability of visual aesthetics of the facility. Conversely, two impacts such as cost and pollution release

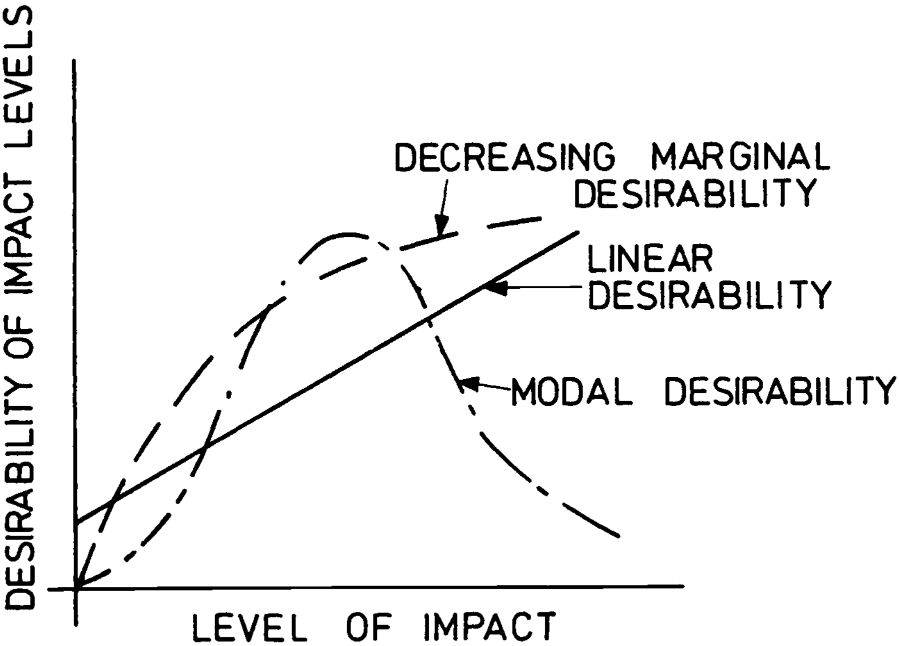

DESIRABILITY OF IMPACT LEVELS
DECREASING MARGINAL DESIRABILITY
LINEAR DESIRABILITY
MODAL DESIRABILITY
LEVEL OF IMPACT

FIGURE 10

may be technically dependent but preferentially independent; the *desirability* of a unit decrease in pollution release might be the same if the facility costs $1.0 million or $10 million. This is a simple but important distinction.

If an objective function specifies linear changes in desirability and independence between the desirabilities of different types of impacts, then the marginal rate of preferential substitution between impacts is constant over all impact levels. This would imply, for example, that if one were willing to increase the facility cost from $10,000 to $10,100 to lower effluent pollution concentrations from 2% to 1%, then one would be equally willing to invoke a cost increase of $10 to $110 to realise a pollution decrease from 10% to 9%. Similarly, if one were willing to increase cost by almost $100 to realise a decrease in pollution of from 1.5% to 1.0%, then one should be willing to increase cost by another $100 (and no more) to realise a further reduction to 0.5%.

Another characteristic of objective functions is whether they reduce evaluation to a single index. In other words, are all impacts aggregated? Methods such as cost-benefit analysis do aggregate, others, such as Bishop's Factor Profile (1972), do not. This represents a philosophical distinction between methods. Although human beings certainty do aggregate in reaching decisions, and politicians or decision-makers must aggregate in any public decision, the issue of dispute is whether or not this may be done explicitly and analytically or only through judgement. Adherents to the former position would say that only in explicitly aggregating can one recognise underlying assumptions and possible biases; adherents to the latter, that the judgemental process of aggregation is so complex that simplified analytical procedures cannot do justice to its full richness and texture. Both arguments have merit. Empirical evidence in experimental psychology (Edwards and Tversky, 1967) would indicate that even the rigorous constructs of rational decision-making represented by utility theory and Bayesian probability does not always perform as well as human judgement. It is difficult to know from historical records whether such theory would have improved decisions made with respect to civil works development (or anything else for that matter). On the other hand, falling back on the sanctity of judgement does open the door to personal biases, and perhaps more importantly to the attempt to grapple intuitively with more impacts than one can remember at any one time. Between these extremes is the idea of aggregating impacts at the sub-objective level in the objectives hierarchy (e.g., aggregating all environmental impacts), and judgementally aggregating across main objectives. This course has the advantage that political decision-makers, while being wary of explicitly weighting impacts against one another -- for example, environmental against financial -- for fear of political repercussions, may be willing to explicitly weight different environmental impacts with respect to one another.

Objective functions also differ in how they treat undertainty in impact predictions. Uncertainty enters predictions in two ways: it may arise from the uncertainty of future conditions such as population density or geophysical phenomena (e.g. floods, earthquakes, tornados), or from an inability to predict (i.e., from lack of knowledge). Inadequate information, e.g., on health effects of radiation, is of the latter type. In terms of the siting decision these two types of uncertainty have identical consequences and are therefore the same. An objective function may either treat uncertainty analytically or leave it as an external for later consideration. In any event, to account adequately for the true net desirability of feasible alternatives, an objective function must explicitly (whether or not analytically) account for uncertainty.

Finally, objective functions differ in the degree to which they are "objective." In the sense we use the term here it means that the analyst's influence on measures of desirability is small. Plan evaluations are always subjective to the degree that they depend on the preferences of people, whether a small group of policy makers or the entire population. However, measures may depend to some extent on non-enumerated interpretations of the analyst, and this is what we take to be lack of objectivity. By this rule elections and many types of market data would be classified as almost purely objective, since little interpretation of the analyst is involved. Colour coding schemes (e.g., Goeller, 1974) and the like are highly non-objective.

4. *Assessment*

All methods of evaluation which would compare favourable and unfavourable impacts of proposed facility sitings to arrive at some ranking rest ultimately on how the assessments of desirability are made. That is, they depend at their foundation on the procedure for collecting desirability data. We have already spoken of attributes as scales along which the impacts of a project can be measured; we must also speak of how to associate desirabilities with those scalings.

All assessment techniques infer desirability from behaviour, it is expressed in the market-place or in replies to an analyst's questions. All assessment techniques make assumptions about the interrelationships of desirability, and then use the structure that derives from those assumptions to draw inferences from empirical data. Very roughly, analysts fall into one of two groups with respect to their philosophy of assessment. This philosophy of the first springs from economic planning theory and views assessment as inference from market data; the second, from sociology and "systems analysis" and views assessment as inference from the direct replies to an interviewer's questions. While these two views might be taken merely as opposite ends of a continuum, it is of interest to look at each in isolation.

A. *Market Approaches*

In a free-enterprise economy it is assumed that the desirability (or utility in an economic sense) of a commodity is reflected directly in the amount of money people are willing to spend for it at the margin. This is a strength of the market-place and the justification for using market prices in evaluating impacts of decisions. For direct impacts of siting, this approach to desirability valuation works well; we have substantial experience with it and understand its pitfalls. Further, the analyst's subjective input is minimised relative to other evaluation techniques, and is relatively easy to discern. Thus there are strong arguments for its use.

Briefly, market approaches first use the set of technological relations to predict impacts along a set of attributes (which need not be monetary units), then associate level of impact on these attributes with monetary values. For example, if an impact attribute were "change of estuary temperature in ^{o}F," one would subsequently associate some monetary cost or benefit with each degree of temperature change. The mapping from attribute to money need not be linear, although in practice it often is. The assignment of monetary units derives from market data either directly or indirectly, and a spectrum of indirect techniques has been developed (e.g., Dorfman, 1965; Layard, 1972; Kendall, 1971).[1] Most of these techniques, however, have been developed to evaluate indirect benefits of a project, while at present techniques for handling indirect *costs* are perhaps insufficient for an adequate accounting (Joskow, 1974; Ross, 1974).

The deficiencies of market approaches, which have often been discussed in the cost-benefit literature (e.g., Dorfman, 1965), are summarised below.

1. Desirabilities of 'non-market' objectives, such as equity, flexibility in future options, and 'balanced' regional growth, cannot be evaluated and thus remain external to the analysis.[2]

---

1. These methods include shadow prices and opportunity costs, compensation costs, willingness-to-pay for or to avoid similar impacts, cost of providing benefits in other ways, and the like.

2. One could argue, of course, that desirability can be expressed in monetary as well as other units; so the degree to which these objectives are met can be associated with monetary desirability. However, this merely transforms the process to one of direct assessment, using money as a scale; it no longer remains a market approach.

2. The use of monetary units implicitly assumes certain interrelationships about desirability, whether they are intended or not -- specifically, linearity over money, independence among impacts, and constant marginal rates of preferential substitution among impacts.

3. Some impacts are very difficult to evaluate because existing market mechanisms are distorted or non-existent (e.g., environmental impacts, health impacts), or because we have no experience with them.

4. Market approaches distort the real desirabilities of impacts toward their market-like facets. The real undesirability of water pollution, for example, may be only partly captured by its economic implications; similar arguments can be applied to reduction of mortality rate, regional development, and other impacts.

B. *Direct Assessment*

Direct approaches go straight to individuals and by means of questionnaires, simple games, and related techniques infer desirability of impacts. These approaches have been developed primarily in the literature of social research and public opinion surveying (e.g., Hansen et al., 1953; Hyman, 1954), and in that of applied decision theory (e.g., Raiffa, 1968).

Opinion sampling is well known, and has well-known pitfalls and biases (Webb et al., 1972); in general these need not be enlarged upon here. Opinion sampling yields qualitative sentiments about the desirabilities of impacts, and most often treats feelings about each type of impact in isolation. (Question: "How would you like to live next to a new highway?" Answer: "Not much.") Often, this means that the results of opinion surveys are difficult to interpret; only in rare cases do they yield quantitative data. The results of opinion surveys do give the analyst or policy maker a good general idea of the sentiments of groups involved, as well as identifying interests (Collins, 1973).

At the other end of the spectrum of direct approaches is the method of "preference assessment" which has been developed in the field of applied decision analysis (e.g., Raiffa, 1968). This approach is oriented towards evoking quantitative statements of preference for impacts and trade-offs among impacts. The method follows from the structure of preference assumed in decision analysis, which in that literature is called a "utility function" (Section VI). Accepting the axioms of preference upon which utility theory is based (Appendix VI.B) leads to an interval scaling of desirability whose mathematical properties can be derived. These properties often allow preferences over a range of impact levels to be estimated by making a small number of measurements.

The procedure for assessing utility functions is based on asking subjects to select preferred alternatives in hypothetical gambles (Appendix VI.A). By presenting hypothetical gambles with multi-attributed outcomes and by varying levels of probability associated with "winning" and "losing", one can have the subject make decisions that force him to implicitly express multi-impact desirability; one can then back-figure preference measures reflected in his answers. Normally, a certain level of redundancy is included in the questioning, and this process is iterated until internally consistent utility functions that the subject retrospectively agrees with are developed.

The strenghts of direct methods vis a vis market approaches is that they allow treatment of impacts with which we have little or no economic experience; that they reflect opinions and feelings which are current (whereas market data are often years old); and that they allow treatment of as yet unrealised impacts, although the whole question of "hypotheticalities" in public or quasi-public decision-making remains a sticky problem.

Opinion surveys and the more quantitative methods of decision analysis are end-points of a spectrum of methods, whose use depends on available time, money, and resources, and on the level of precision required. The question resolves to one of investment in public sampling vs. error in resulting quantifications of desirability. The latter end of that spectrum consists of methods that bring out quantitative trade-offs among the desirabilities of impacts; the data one receives from this end of the spectrum are much more useful than those from the other end, but cost more.

Several important deficiencies of direct approaches are listed below.

1. The ordering and even the wording of questions introduces bias errors of whose magnitude and direction the analyst is ignorant.

2. Subjects may have preferences for impacts but be unable or unwilling to verbalise them.

3. Even if, after great introspection, a subject can verbalise his preferences, are these the same as would be inferred from his behaviour (i.e., in action) and how could you ever find out? If it is not, which is more proper? Clearly one would be measuring something different other than what is measured by market approaches.

4. Cost constrains the number of individuals interviewed and the depth of the interviews. This leads to larger "estimation errors" than market approaches which generally have larger data bases.

5. Assessment techniques involve hypothetical gambles and therefore depend not only on subjective preference but on subjective probability as well.

6. Non-naive subjects sometimes deliberately mislead interviewers in the hope of biasing decisions toward their true preferences (i.e., "gamesmanship," or what Swain (personal communication) calls the "garden path effect").

C. *Combined Approaches*

There is no reason why market and direct approaches cannot be combined for a better description of desirability than either approach leads to in itself. This is generally not done because analysts approach problems with a pre-chosen decision methodology, carrying with it a philosophy of assessment.

While work is needed to develop a combined approach, such an approach might use market techniques to measure economic impacts or impacts that are easily and justifiably treated with market data, and direct assessment to measure non-market impacts (and those which are difficult to measure behaviourally, such as mortality rate). Sets of assessments could overlap, and could be calibrated with respect to each other to reduce bias errors. A second approach would be based as this one, but use market date as prior information in the Bayesian sense, and modify those data by direct assessments in the normal Bayesian scheme of updating (Baecher, 1975).

Table 1

(after Stevens, 1959)

Scale	Empirical Operations	Group Structure	Measure of Location	Dispersion
Nominal	determination of equality	permutation group $x' = f(x)$ where $f(x)$ is any one-to-one substitution	mode	information, H
Ordinal	determination of greater or less than	isotonic group $x' = f(x)$ where $f(x)$ is any monotonically increasing function	median	percentiles
Interval	determination of the equity of intervals or differences	linear group $x' = ax + b$ $a > 0$	arithmetic mean	standard deviation
Ratio	determination of equity of ratios	similarity group $x' = cx$ $c > 0$	geometric mean harmonic mean	percent variation

APPENDIX

Measurement Theory

One assigns numbers and symbols to events and objects because mathematical relationships among properly defined numbers and symbols have been extensively studied and are well known. Since some of these relationships may be shared by the events and objects, one may by analogy infer properties of the events and objects that have not been observed or are not immediately obvious. However, one must be explicit about relationships among the events and objects, because numbers and symbols may be related in ways in which the events and objects are not (Ackoff, 1962).

The relationships one assumes to hold between the events and objects one assigns numbers to are implicit in the scale used. The following four scales are generally recognised.[1]

1. Nominal Scales group elements into classes; for example, a facility site might be either inland or coastal.

2. Ordinal Scales rank elements with respect to some dyadic relationship (i.e., "greater or less than" relationships). The Mercalli scale of earthquake intensity is an ordinal scale.

3. Interval Scales introduce a unit of measurement; distances between elements on the scale represent distances between them in some relationship defined over them. The Centigrade temperature scale is an example.

4. Ratio Scales introduce the property of absolute zero in addition to interval properties; ratios of scale values represent ratios in the relationship defined over the elements. Money is a ratio scale.

The scale to which events or objects are measured also defines permissible mathematical and statistical operations on the resulting measurements (Table 1). Because the scale specifies allowable operations, the operations required by an evaluation methodology dictate the level of scaling required. Simple comparison of deterministic impacts requires only ordinal scaling (e.g., indifference curves -- Section VII); (e.g., von Neumann-Morgenstern Utility -- Section VII); ratios of desired to adverse impacts require ratio scaling (e.g., cost-benefit analysis -- Section V). Applying inadmissible operations to measurements result in numbers whose relationships to one another have no meaning. For example, if different alternatives have impacts against some objective whose desirability we can ordinally scale (best, second best, ..., worst), and if we assign the numbers 1,2,...,n to those desirabilities, then we cannot add the desirabilities together nor weight them to form an aggregate average.

IV. *Cost Benefit*

Ever since Dupuit observed that more general benefits accrue to society than are manifested in revenues, decision-makers have been searching for techniques that can include all of these in one analysis. Perhaps the most-used technique is cost-benefit analysis. Here, a project is analysed by summing economic benefits to all of society and comparing them with economic costs; if the former exceed the latter, then the project is either deemed favourable for investment or ranked against alternatives. Cost-benefit has been subject to debate and refinement for decades. The purpose here

---

1. Stevens (1959) and Stevens and Galanter (1958) suggest others, but they are primarily of theoretical interest.

is not to present the spectrum of opinion, but to review some basic or implicit assumptions of the technique, to discuss the ease of applying it for site evaluation, and to compare it to other methods of analysis.[1]

During the New Deal era, cost-benefit analysis was adopted in the United States as a tool to evaluate public works programs. The returns on these projects were often insufficient to interest private investment, but were attractive to the government because total benefits often exceeded costs. The Flood Control Act of 1936 institutionalised the use of cost-benefit analysis, which has remained the primary tool for evaluating public works programs ever since. This Act set the important precedent for U.S. government policy that benefits "to whomsoever they accrue" should exceed costs, and did not require an enumeration of the recipients. Since the Act, the U.S. government has made major efforts to incorporate modifications and extensions into the general procedure (see U.S. studies of 1965, 1971), and cost-benefit techniques have been applied to decisions in such disparate fields as public health, outdoor recreation, and defense, and in both the public and private sectors (Dorfman, 1965).

In cost-benefit analysis the only criterion of decision is economic efficiency. This criterion has traditionally been taken either to be the ratio

$$B/C = \frac{\Sigma b_i}{\Sigma c_i} > 1.00, \qquad (3)$$

or the difference

$$B\text{-}C = \Sigma b_i - \Sigma c_i > 0, \qquad (4)$$

where the b_i's and the c_i's are benefits and costs, retrospectively, expressed in monetary terms.

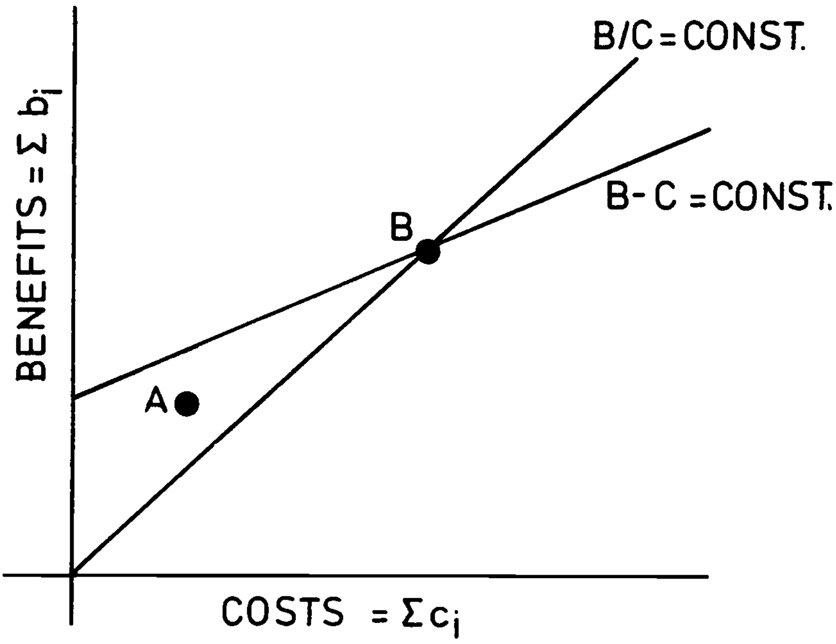
Benefits are commonly separated into direct and indirect. The former include the immediate products or services of the project, often expressed by direct revenues; the latter include all other benefits accruing from the proposed project, such as increases in regional economic development, flood protection, etc. Costs can be similarly divided, and again the summation includes both.

When used to generate an ordinal ranking of plans, the alternative with the largest benefit to cost ratio or benefit less cost difference is preferred, followed by the one with the next-highest, and so on. In public expenditure practice, however, cost-benefit analysis often serves as an admissibility test in which all alternatives with a $B/C < 1.0$ are screened out and decisions among those which remain are made on other bases (Sewell, 1973). When an ordinal ranking is generated, the benefit/cost ratio and benefit-cost difference can lead to different orderings of alternatives, as the ratio criterion favours low-cost alternatives (disregarding economies of scale) while the difference criterion favours high-cost ones (Figure 11). Given several projects with constant total budget, the ratio criterion can easily be shown to maximise net return; while for any one project with no cost constraint, the difference criterion obviously maximises net benefit.

Siting decisions are different from the usual budget allocation problem in that the value of benefits is usually considered to be independent of the site considered. Therefore, after the decision has been made to build the facility, the problem is more

---

1. More detailed reviews and discussions of cost-benefit theory and its problems are given in Prest and Turvey (1955), Mishan (1971), Maass et al. (1962), Marglin (1967), Eckstein (1958), and UNIDO (1972).



(SITE A HAS A GREATER BENEFIT/COST RATIO BUT LESSER BENEFIT-COST DIFFERENCE THAN SITE B)

FIGURE 11

nearly a cost minimisation problem than a true cost-benefit problem. Perhaps this can best be characterised as a cost-effectiveness approach.

The primary advantages of the cost-benefit technique relative to other decision tools are:

1) It is conceptually simple and readily understandable, and decision-makers have experience in using it;

2) It has a basis in general welfare theory, although it is normally used more pragmatically (Broadway, 1974);

3) It reduces multi-dimensional impacts to one scalar index for easy comparison of alternatives;

4) It attempts to be objective, limiting the analyst's influence on the results.

The disadvantages are:

1) The use of monetary units for all impacts places restrictive assumptions on the preference structure and does not allow inclusion of more than one group's values or more than one averaging of "society's" values;

2) It does not include many social objectives;

3) It lacks a satisfactory way of treating uncertainties in impact predictions;

4) By reducing impacts to monetary units, it leads to market-like approaches to evaluation, which often involve complex schemes not fully capturing the true desirability of impacts.

In cost-benefit analysis, all impacts are expressed in monetary units. Two restrictive and probably unrealistic assumptions about the preference structure result:

1) Desirability is a linear function of impact level for each impact.

2) The desirability of any impact level is independent of the levels of other impacts.

These implicit assumptions result in restrictions on the marginal rate of substitution between impacts (i.e., it is assumed constant).

The disadvantages listed as Nos. 2 and 4 deal with what are known in cost-benefit analysis as *externalities*. These are impacts that, while important, cannot be included in the decision analysis in ways which adequately reflect their true importance. Some of these are noise, health and safety impacts, environmental degradation, and social disruption. To the extent that externalities relate to important objectives, cost-benefit analysis is incomplete and can be only one of several factors in reaching a final decision.

Economists have been clever in including in the cost-benefit framework impacts that would seem at first appearance to be inexpressible in monetary units (noise, for example; Heath, 1971). Often, however, such impacts are treated by establishing legal standards or constraints that must be met in decision-making rather than treating the impacts as merely another variable. This suggestion has been made by Joskow (1974), for example, with respect to siting nuclear facilities. The approach is not at all satisfactory, because it simply transfers responsibility for decisions to another place, in this case to regulatory agencies. If they are making their standard setting decisions with the same cost-benefit methodology (see, e.g., Majone, 1974) we are still left with the problem.

1. *Equity*

Implicit in cost-benefit analysis is a disregard for the distribution of impacts. An alternative that greatly benefits a few people while adversely affecting many or even most, is perfectly admissible as long as its benefits to society as a whole exceed its costs. In siting decisions, these questions of equity refer to the distribution of effects both over the strata of society and over spatial groups.

There have been many attempts to include questions of equity in the cost-benefit framework. A common approach is to list efficiency calculations alongside equity (and other "non-scientific" criteria) in presenting alternatives to decision-makers, who are then called on to make subjective comparisons. This approach was used by the Roskill Commission (1970) on siting the Third London Airport, and was recommended by the Water Resources Council (1971) for U.S. Government projects. By including equity considerations in this manner, cost-benefit analysis becomes similar to some of the matrix methods discussed in the next section.

Marglin (1962) suggests the use of constraints on costs and benefits accruing to groups. The problems with this method, however, are that constraints must be chosen arbitrarily, and that there is no provision for trade-offs between efficiency and equity (Weisbrod, 1968). A second method is to apply weighting factors to benefits and costs for each group, and then take a weighted sum over all groups. Values of the first weights would correspond to values that groups themselves attach to changes in particular impacts, and the second set of weights would correspond to the importance of each group having its preferences satisfied (i.e., political weights). Weisbrod has suggested that the political weights might be inferred from past government decisions.[1] Weights of this type assume independence among the groups.

On the other hand, many applications of cost-benefit analysis simply ignore equity. Justifications of this are usually taken to be (Layard, 1971):

1. The so-called "Hicks-Kaldor criterion", which says that one should be concerned only that beneficiaries *could* compensate losers even if in reality they don't; a concept often extended by the concpet that adverse distributional effects can be undone by purely redistributional projects;

2. The impropriety of undertaking interpersonal comparisons of the marginal value of benefits and costs;

3. A multiplicity of projects will tend to even out distributional effects.

2. *Uncertainty*

Siting decisions involve uncertainties, with respect not only to health and safety impacts, but also to a range of social, environmental, and even monetary costs; and any rational decision process must provide a means of accounting for them. Uncertainties result from (a) random events, such as weather conditions, future population

---

1. This method circumvents a value judgement by the analyst by using the value judgement of politicians. The interesting objection has been made by Layard (1972) that if past decisions were consistent and rational, why not continue in the same process; and if they were not, why pretend that they were?

Sector and Instrumental Objective	Measure	Differences (£ m.) from Cublington			
		Cublington	Foulness & Luton	Nuthampstead	Thurleigh
PRODUCERS OPERATORS					
Air and Surface Transport					
British Airports Authority					
Airport Construction	£	0	+32	- 4	-18
Operating Costs	£	0	-22	- 5	-15
Airline Operators					
Meteorology	£	0	- 5	- 3	- 4
Airspace Movements	£	0	5	31	26
Accident Hazards	£	0	2	0	0
Highway Authorities					
Capital Costs	£	0	4	4	5
Public Transport Authority					
Capital Costs	£	0	+23	+ 9	- 3
DISPLACED OR AFFECTED PRODUCERS					
Defence	£	0	-29	-24	+32
Public Scientific Establishments	£	0	- 1	+20	+26
Private Airfields	£	0	- 7	+ 6	+ 8
Schools, Hospitals & Public Authority Buildings	£	0	- 2	+ 4	+ 2
Agriculture	£	0	+ 4	+ 9	+ 3
Commerce and Industry	£	0	+ 2	+ 1	+ 2
Producers: Total:	£	0	+ 6	+48	+64
CONSUMERS					
TRAVELLERS AND FREIGHT SHIPPERS					
Passengers					
(a) On Surface: British residents	£	0	131	28	17
: Foreign residents	£	0	36	7	5
(b) In the Air (included in 1.2)					
Freight Shippers	£	0	14	5	1
Other Travellers (included in 2.1)					
DISPLACED OR AFFECTED CONSUMERS					
Residents Displaced	£	0	-11	- 3	- 5
Residents Not Displaced					
-Noise: 55 NNI+	£	0	0	2	3
50-55 NNI	£	0	- 1	- 1	0
45-50 NNI	£	0	- 1	+ 3	- 4
40-44 NNI	£	0	- 1	+19	- 5
35-40 NNI	£	0	- 2	+27	0
-Recreation	£	0	-13	- 6	- 6
RATEPAYERS, TAXPAYERS AND GENERAL PUBLIC		-	-	-	-
Consumers: Total:	£	0	+152	+81	+ 6
Overall Total:	£	0	+158	+129	+70

Figure 12

Balance Sheet of Development
(after Lichfield, 1971)

levels, and equipment failures, and (b) lack of information on long-term consequences. As we have already argued, these should be treated similarly.

A satisfactory method of handling uncertainty in cost-benefit analysis has yet to be developed (Dorfman, 1962), although several methods have been explored and applied. Among these are: using expected values of impacts, trying to assess certainty equivalents, and using discount factors.

The most straightforward approach is to use an impact's expected value in cost-benefit analysis. This corresponds to linear preferences for money in uncertain situations; while expected value may be legitimate over small uncertainty ranges, it is likely to be legitimate over large ones. Thus expected monetary value is not the same as expected desirability, and we have the intuitive contradiction that distributions of possible impact values are equally desirable as long as their mean values are the same. The second approach is to specify a certain impact for which one would be indifferent to the choice between it and the uncertain impact. Much of the "risk evaluation" work in nuclear power uses this approach (Otway et at., 1971; Starr, 1970). Often, however, certainty equivalents are determined on an *ad hoc* basis, and cannot be back-figured using utility functions and economic data. A critical discussion of this approach is found in Dorfman (1962). A common heuristic technique is to discount the expected value of impacts by some measure of the uncertainty; a typical factor is $(1 + k\sigma)^{-1}$ where k is a positive constant and σ is the standard deviation.

The drawbacks of all three methods are that they are simply rules-of-thumb (Eckstein, 1961; Dorfman, 1962) with no sound theoretical basis.

V. *Matrix Approaches*

Given the multi-attribute nature of impacts from siting large facilities and what is seen to be an inherent non-comparability of impacts of different types, several methods of project evaluation have been developed which list impacts separately in a table or matrix (Figure 12). These methods hope to circumvent apparent non-comparabilities by allowing the decision-maker to choose a best decision alternative judgementally after reviewing the spectrum of differeing impacts.

While several "matrix" approaches have been developed, they spring from the same philosophy; impacts against different types of objectives are inherently non-comparable; it is true that people do make decisions that require implicit trading-off of one type of impact for another, but schemes to analyse such trade-offs quantitatively invariably stumble over the necessary simplifying assumptions. While trade-off relations might be developed on subjectivist theory, as in utility theory, the analysis cannot do justice to the full complexity of judgemental decision-making, and some impacts of large facilities simply bar quantification.

In this section we will present four groups of matrix techniques which embody a range of those proposed, and conclude by summarising the advantages and limitations of non-aggregating approaches to siting.

1. *Lichfield's Planning Balance Sheet*

Lichfield's (1968, 1971) planning balance sheet method is an outgrowth of cost-benefit analysis which received renewed attention in the wake of controversy over the Roskill Commission's analysis of sites for the Third London Airport. This method attempts to separate from one another both impacts considered inherently non-comparable, and those against different groups within society. Typically, a planning balance sheet might look like that schematically illustrated in Figure 12, in which monetary units are used for impacts that may be readily so quantified, and non-monetary units for the remainder. If an impact is judged to be non-quantifiable numerically it is assigned qualitative descriptions. Impacts expressed monetarily are aggregated as in normal cost-benefit analysis, and a decision is made judgementally by weighting the net monetary cost of benefit against the spectrum of other impacts and their distribution across groups.

Goal Description		α			β			γ			δ	
Relative Weight		2			3			5			4	
Incidence	Relative Weight	Costs	Ben.	Relative Weight	Costs	Ben.	Relative Weight	Costs	Ben.	Relative Weight	Costs	Ben.
Group a	1	A	D	5	E	-	1	-	N	1	Q	R
Group b	3	H	-	4	-	R	2	-	-	2	S	T
Group c	1	L	J	3	-	S	3	M	-	1	V	W
Group d	2	-	-	2	T	-	4	-	-	2	-	-
Group e	1	-	K	1	-	U	5	-	P	1	-	-
		$\sum$	$\sum$					$\sum$	$\sum$			

Figure 13

The Goals-Achievement Matrix
(after Hill, 1973)

The advantage of Lichfield's method over traditional cost-benefit is that it explicitly enumerates impacts that seem "unmeasurable" (and thus are not normally included) and specifies the distribution of impacts over affected groups. However, it gives no guidance to how these might be incorporated in a decision, other than that impacts on groups might be weighted to account for equity considerations.

2. *Goals-Achievement Matrix*

The "goals-achievement" approach developed by Hill (1973) is perhaps the most widely publicised of the various matrix techniques. Hill uses the term *goal* in precisely the same way as we have used the term *uppermost objective;* sub-objectives, lower in the hierarchy, he merely called *objectives*.

The essence of the goals-achievement approach is to establish separate accounts for impacts generated by contending sites and technologies as they bear against each important goal and each of several groups within society. Achievements toward each goal and impacts against each group are given weights on judgemental bases, and those levels of goal achievement (multiplied by their appropriate weight) which are in commensurable units are combined, leaving a reduced but still multi-dimensional array to be reviewed in reaching a final decision. The method is one step closer to aggregation than simple impact display tables, but again breaks down when the number of unaggregated impacts becomes too large for intuitive treatment.

The procedure for generating a goals-achievement matrix is the following. First, each goal of importance is identified, and attributes with which to measure achievements against each is selected. If a quantitative index cannot be associated with each goal, a qualitative description of predicted impact is substituted. Second, weights are judgementally assigned to each goal on the basis of its importance; each population group affected by the proposed project is identified, and the importance of impacts on each group with respect to each goal is weighted. Finally, these are arranged in matrix format as shown in Figure 13 (in which capital letters represent costs and benefits, in a generic sense, accruing to each affected group). Cost and benefits with respect to each goal must be in similar units, and if these are quantified predictions, the weighted sum over all affected groups is "meaningful", then a "grand cost-benefit summation" is possible.

The goals-achievement matrix, like other matrix approaches, includes no analytical way of treating uncertainty. Although Hill readily admits (1973, p. 27) that "uncertainty concerning anticipated consequences is best treated by probability formulation", the most that is currently done is to include ranges of possible impacts rather than point estimates. "In general, allowance for uncertainty should be made indirectly by use of conservative estimates, requirement of safety margins, continual feedback and adjustment and a risk component in the discount rate" (1973, p. 28). This does not seem satisfactory.

To this point the goals-achievement matrix is only a vehicle for displaying predicted impacts of site and facility technology alternatives. Given this listing, how is a decision or ranking of alternatives made? Hill suggests three techniques of varying levels of aggregation. The simplest is just to let the decision-maker review the matrix and arrive judgementally at a decision; at this level the method is primarily bookkeeping. The next level is to aggregate impacts using the weightings assigned to goal achievement and group impact, but here the method adopts those very inadequacies it was developed to mitigate. According to Hill (p. 37), "the combined weight of the objectives and their incidence is assigned to the measures of achievement of the objectives. The weighted indices of goals-achievement are then summed and the preferred plan among the alternatives compared is that with the largest index". Clearly this approach differs little from traditional cost-benefit analysis except that units other than money may be specified explicitly rather than being hidden in specified monetary values. The explicit weighting of impacts on groups is similar to Lichfield's planning balance sheet and Weisbrod's (1968) suggestions for traditional cost-benefit analysis.

The central problem with aggregation of impacts in this way is that it assumes interrelationships in the objective function (i.e., the desirability of impacts relative to one another) that may not be reflected in reality. Namely, it assumes that the degree to which we should desire a certain level of an impact is independent of the levels of all other impacts, and of the level of that impact against that same goal relative to other groups; and is a linear function of absolute level with a defined zero point. It is not at all clear that these even approximate valid assumptions; and so the goals-achievement approach contributes little to over-coming the limitations of cost-benefit assumptions.

Hill goes on to say that although not every impact may be scaled on cardinal indices, the goals-achievement method may be modified to handle ordinally scaled impacts. His proposed method would assign the value +1, 0, or -1 to each impact on each group, depending on whether it enhanced, left unchanged, or detracted from goal achievement. These *ordinal* values would be combined by multiplying each by both the goal and the group weight and summing to determine a final aggregate index of goal attainment. This is blatantly erroneous: if impact data are specified to an ordinal scale they do not allow multiplication and addition, so the final index is meaningless.

Hill's final proposal is based on Ackoff's (1962) notion of transformation functions which map one impact scale onto another, and approaches the concept of measurable utility which is treated in Section VI. Hill suggests that impacts that are measurable to either an interval or ratio scale be transformed onto one common scale through some (not necessarily linear) transformation. In the two-impact case this would mean expressing levels of one impact in units of the other. As the correspondence between increments of impacts is not necessarily constant over the ranges of those impacts, these transformations might not be linear. In the multi-impact case the easiest proposition might be to scale all impacts in terms of a single impact, perhaps money. In this case, Hill's proposal once again reduces to a form of cost-benefit analysis, except that non-linearities in the evaluation of impact levels would be allowed. This does not circumvent other assumptions of independence or allow one to treat analytically impacts defined to less than an interval scale, as discussed previously. Given that this approach attempts to express quantitatively trade-offs between the desirability of different impacts and non-linearities in the desirability of levels of one impact, there seems little reason not to go over entirely to a utility analysis, which makes few additional assumptions and is more theoretically sound.

3. *Environmental Impact Matrix*

Leopold et al. (1971) of the U.S. Geological Survey have presented what they call an "environmental impact matrix" for use in compiling environmental impact statements as required by the Environmental Policy Act of 1969. This technique is primarily intended to provide a uniform procedure for coalescing impacts and presenting them, rather than being a decision-making tool in itself. As the authors state their intention,

> "The heart of the system is a matrix which is general enough to be used as a reference checklist or a reminder of the full range of actions and impacts on the environment that may relate to proposed actions".

Their hope is to provide "a system for the analysis and numerical weighting of probable impacts" which would "not produce an overall quantitative rating but portrays many value judgements".

In essence the environmental impact matrix is intended to be a tabular summary of project impacts which would accompany environmental impact statements. But as this method attempts to scale impacts, and as some workers have attempted to use it as a decision tool, a few remarks are in order.

	Industrial sites and buildings	Highways and bridges	Transmission lines	Blasting and drilling	Surface excavation	Mineral processing	Trucking	Emplacement of tailings	Spills and leaks
Water quality					2/2	1/1		2/2	1/4
Atmospheric quality						2/3			
Erosion		2/2			1/1			2/2	
Deposition, Sedimentation		2/2			2/2			2/2	
Shrubs					1/1				
Grasses					1/1				
Aquatic Plants					2/2			2/3	1/4
Fish					2/2			2/2	1/4
Camping and hiking					2/4				
Scenic views and vistas	2/3	2/1	2/3		3/3		2/1	3/3	
Wilderness qualities	4/4	4/4	2/2	1/1	3/3	2/5	3/5	3/5	
Rare and unique species		2/5		5/10	2/4	5/10	5/10		
Health and safety							3/3		

Figure 14

Environmental Assessment Matrix
(after Leopold et al., 1971)

The matrix is constructed by listing aspects of a proposed alternative that might produce impacts along one axis, and types of impacts along the other (Figure 14). In each resulting square of the matrix with which significant impacts are associated, two numerical entries are made: the upper, a measure scaled on the integer range (1,10) indicating the magnitude of impact; and the lower, again a measure on the integer range (1,10), indicating importance of impacts. Although these numbers are assessed judgementally, to the extent possible they "should be ... based on factual data rather than preference". Although the authors are not specific about how this should be done, they suggest that such a quantification "discourages purely subjective opinion". This does not seem immediately true; more likely, such quantification requires the analyst to be more honest in his subjective evaluation of impacts, which will be uncompromisingly stated in his report and open for direct questioning -- as with any quantification. The environmental impact matrix provides no mechanism for treating uncertainty, and the authors make it very clear that one should not try to compare impacts from square to square on the same matrix.

As a summary chart this method is not without merit, except that quantification as presented here can easily be misinterpreted. Some workers (e.g., Beer, 1974) have attempted to coalesce these impact measures by forming the weighted sum of matrix entries (the very thing cautioned against in Leopold et al., 1971), which not only presumes the assumptions of additive desirability but takes impact indices to be intervally rather than ordinally scaled.

4. *Bishop's Factor Profile*

Bishop's "factor profile" (1972) is in essence a graphical technique for displaying project impacts. However, it has received some mention as a decision-making tool (e.g., Fischer and Ahmed, 1974) and so will be briefly reviewed. A typical factor profile is shown in Figure 15. In this profile each non-financial impact is scaled on an (-100, +100) interval range on the basis of its relative desirability, -100 being the least desirable and +100 the most desirable of the impacts of contending alternatives against that goal. A decision is reached via a four-step procedure:

1) the economic impact of each alternative is determined in benefit to cost ratios,

2) factor profiles are constructed for each alternative,

3) dominated alternatives on both the factor profile and benefit/cost ratio are eliminated,

4) pair comparisons are made on the remainder to assess relative desirability (judgementally), and an ordinal ranking is thus generated.

Factor profiles are more a graphical display device than a decision tool, thus offering little that Lichfield's balance sheet does not. Although Bishop does not extend factor profiles to the separation of group impact, this could be accomplished with minor alteration. The assumption of interval scaling seems more restrictive than necessary, as ordinal scaling is all that is required.

5. *Advantages and Disadvantages of Matrix Methods*

The advantages of the matrix methods reflect the disadvantages of cost-benefit analysis that they were designed to overcome. Their primary advantage is that they allow the explicit inclusion on non-efficiency objectives in an analysis, although they do not indicate how one should trade off achievement of economic and non-economic objectives. However, many proponents of matrix methods would say that such trade-offs are inherently non-quantifiable and thus can be made only in a purely judgemental way. This works satisfactorily when the number of non-aggregable impacts is small, but not when it is large: still then there is a danger of biasing a decision toward economic objectives

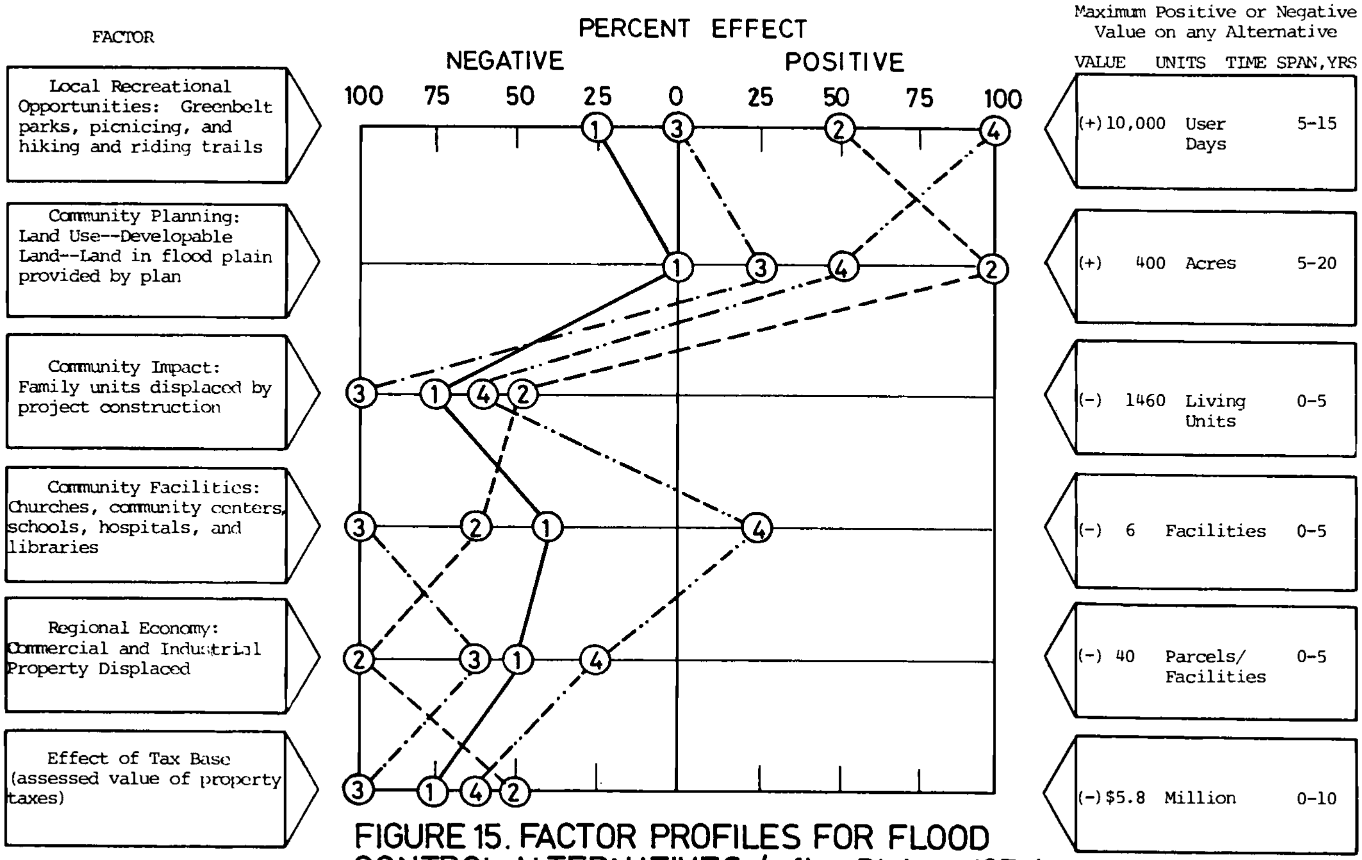


FIGURE 15. FACTOR PROFILES FOR FLOOD CONTROL ALTERNATIVES (after Bishop, 1972)

as the spectrum of impacts is so large that a fuller integration is conceptually difficult.

Secondary advantages of matrix methods are that they are good vehicles for presenting impacts to decision-makers, and that they do not require quantification of certain impacts, such as aesthetic ones, that are difficult to scale.

The central disadvantage of matrix techniques is that they do not tell one how a decision should be made, and when secondary procedures are used for considering the totality of impacts they often lead to misinterpretations. In particular, the schemes that have been used to aggregate matrix entries usually assume that there is independence among the desirabilities of impacts, and that one may perform mathematical operations with what are often ordinally scaled quantities.

VI *Preference Theories*

The methods discussed so far assign desirability to impacts and thus generate objective functions based on economic impact or simple weighting schemes. Although some of these methods carefully scale relative desirabilities of levels of simple impacts, none adequately accounts for interaction among impacts. That is, they assume that marginal changes in the desirability of levels of one impact do not depend on levels of associated impacts; these desirabilities are independent.[1]

There does not exist a set of methodologies, however, in which the desirabilities of multi-attribute impacts are rigorously handled, including interdependencies among impacts. These methodologies are based on a set of simple axioms of preference, and from this axiomatic foundation mathematical properties of multi-attribute objective functions are derived. In this way interrelationships are explicitly stated, in contrast to previously discussed methods in which they were implicit and therefore often neglected.

These methods are explicitly based on the tenet that desirability of impacts derives from subjective preferences rather than so-called "objective" criteria, citing the failure of general welfare theory to provide that objective valuation.

We will discuss the theoretical foundations of three levels of axiomatically based preference functions, and then turn to their application in siting and a discussion of their advantages and disadvantages relative to other methodologies.

If one assumes that a preference ordering can be assigned for any pair of impacts or impact levels (that is, if for any pair of impacts A and B, either A is preferable to B, or B is preferable to A, or A and B are equally preferable), then a preference ordering over an entire set of impacts can be constructed. Further, if the preferability of pairs of impact levels can be assessed relative to other pairs of impact levels (that is, if given two types of impacts X and Y and two levels of each impact X_i, X_j and Y_i, Y_j, the relative preferability of the pairs (X_i,Y_j), (X_j,Y_i) can be assessed(, then a family of "indifference curves" can be generated (Figure 16) with the property that any two pairs of impact levels on the same indifference curve should be equally preferable (e.g., (X_i,Y_j), (X_j,Y_i)). Applying similar arguments, one can generate indifference surfaces in higher-order spaces (Fishburn, 1970) and thus an ordinally scaled objective function for evaluating the desirability of specified sets of impact levels.

The important thing to note here is the indifference surfaces are ordinally scaled; the normal operations of multiplication and addition are not defined over them, and

---

1. An argument could be made that cost-benefit analysis circumvents this interaction, because in economic efficienty terms the desirabilities of impacts <u>are</u> independent; but this is a narrow case and leads to the common objection that we should make evaluations on broader grounds.

INDIFFERENCE CURVES

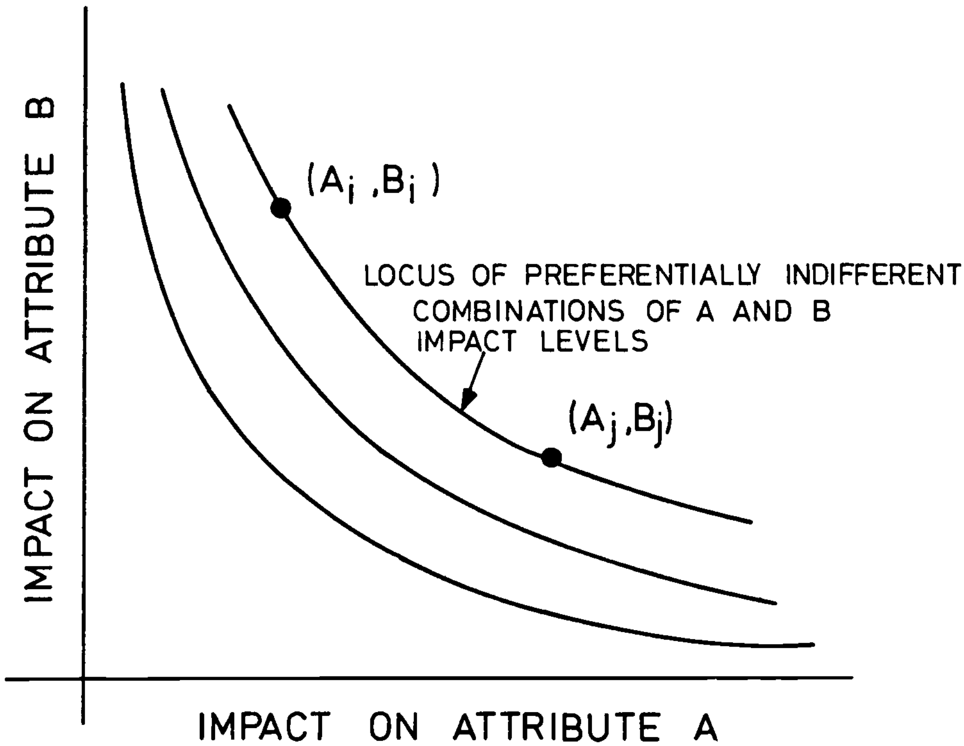


FIGURE 16

common procedures of reducing the work of assessment and evaluation are not allowed. To assess a set of indifference curves requires individual assessment of the relative preferability of each point in the multi-dimensional space and entails substantial effort -- too much, in fact, to be reasonable for more than, say, three or so impact attributes. Further, there is no rigorous way to include uncertainty in the analysis, again because the ordinal scaling does not allow arithmetical operations.

Despite these drawbacks in implementation, indifference surfaces have been used in siting and project evaluation, most notably in the work of Major (1974) and MacCrimmon (1968). MacCrimmon and Toda (1969) have also described a procedure for obtaining indifference surfaces. An advantage of indifference surfaces is that the additional assumptions necessary to develop integrally scaled functions need not be introduced, yet varying marginal rates of preferential substitution among impacts can be represented.

1. *Value and Utility Functions*

If an expanded set of axioms on preferability between impacts is introduced, integrally scaled preference functions can be derived. This results in a function similar to indifference surfaces but for which each surface represents a contour of preference which can be assigned a numerical value, and for which the differences between these numerical measures carry meaning. This allows the mathematical operations defined on integral scales to be performed on the preference function; such functions are generally called "value functions".

By increasing the set of axioms (Appendix VI.B) and by modifying the procedures of assessment, value functions can be expanded to apply to cases in which impact levels are uncertain but can be described by probability distributions. The latter function has become widely known as *utility*, or sometimes *measurable utility*, in differentiation to the classical concept of utility in economics.

Smith (1956) has presented a historical summary of utility theory. Although beginnings of the theory can be traced as far back as Daniel Bernoulli, it has seen the bulk of its development in the past 25 years. A rigorous treatment of the foundations can be found in Fishburn's writings (e.g., 1964, 1970).

2. *The Utility-Based Decision Model*

Given the axioms of utility theory, an optimum decision is that which leads to a maximisation of expected utility (Pratt, Raiffa and Schaifer, 1965). In the notation introduced in Section III, the set of decision alternatives leading to the most preferred set of impact levels is that which maximises

$$E[u] = \int \int u(\underline{x},\underline{\Theta} | s,q) \; f(\underline{x},\underline{\Theta}|s,q) \; d\underline{x}d\underline{\Theta}, \qquad (5)$$

where $u(\underline{x},\underline{\Theta}|s,q)$ is the utility function. Although this is conceputally straightforward, in practice the process is made difficult because the utility function itself can become complicated unless certain properties of the structure of preference are shown to apply, and because assessment of utility functions is an involved task. Given also that utility theory is based on subjective preference, the question of whose preference structure to use is more explicit here than in other methods, even though one can forcefully argue that none of the methods are truly "objective"; thus "whose objective function to use" is always a problem.

3. *Form of the Utility Function*

Unless certain restrictive properties of the interdependence of preference over different types of impacts can be assumed to apply in a particular case, the mathematical form of the utility function can be quite complicated and even approach intractability. Keeney (1972) has reviewed forms of multi-attribute utility functions, and has shown that two "independence properties" are of critical importance in establishing the appropriate form. These are called *value independence* and *utility independence*. Value independence is the more restrictive of the two and is a sufficient condition for utility independence; utility independence is only a necessary condition for value independence.

Value independence is the property that preferences for gambles depend only on the marginal (i.e., single variable) probability distributions of impacts and not on their joint (i.e., multivariate) probability distributions.

Utility independence is the property that preferences for gambles involving uncertainties in one impact, conditioned on known values of the other impacts, do not depend on what those other values are.

We will not dwell on definitions of these properties, for they are presented elsewhere (e.g., Keeney, 1973). The important thing to note is that only if value independence holds is the simple additive form of the multi-attribute utility function appropriate:

$$u(\underline{x}) = \sum_{i=1}^{n} k_i u_i (x_i) \; . \qquad (6)$$

If utility independence holds sufficiently often, then either the additive form or the multiplicative form,

$$1 + ku(\underline{x}) = \prod_{i=1}^{n} \left[1 + kk_i u_i (x_i)\right] \; , \qquad (7)$$

[1]

may be appropriate, depending on whether value independence also holds. Again, *unless one of these properties holds, the additive or multiplicative forms of the multi-attributed utility function are not applicable.*

This greatly increases the difficulty of assessment and, if the decision structure contains continuous variables, also reduces the mathematical tractabiltiy of optimisation.

In the siting and environmental impact literature, additive forms of the utility function are widely used and only infrequently justified by attempts to demonstrate value independence -- or at times even to mention it. The whole set of decision methodologies which use rating scales for individual impacts and a weighted sum for aggregation are forms of additive utility and incorrect in preferential terms unless the restrictive condition of value independence holds.

A problem with applying utility theory to siting decisions is assessing utility functions. This can be a long process and requires some degree of familiarity with the technique by individuals whose preferences are being assessed. Further, a satisfactory procedure for measuring group utility functions, when they are to be used, has yet to be developed. These drawbacks were discussed in Section III.

4. *Application*

While cost benefit and matrix methods have been used extensively in plan evaluation and siting, utility models have been used only infrequently. An initial application of utility to siting public facilities was made by de Neufville and Keeney (1974) on

---

1. Both k and k_i are constants with the properties $\sum_{i=1}^{n} k_i = 1$ in the additive form, and $\sum_{i=1}^{n} k_i \neq 1$ in the multiplicative form.

the problem of siting the new Mexico City Airport. In that work the authors used an impact set consisting of six objectives and attributes, of which three dealt with cpst and service and three with social/environmental effects: safety, social disruption (as measured by the number of people displaced by construction), and noise pollution. In the final analysis, however, the problem was seen to be an innately political one dealing with phasing levels of commitment to opposing sites.

An attempt to apply utility models with a limited set of objectives to power plant siting in New England was made by Gros (1974), who also addressed the problem of differing interest groups having different utility functions. However, in neither the de Neufville-Keeney nor the Gros study were utility functions directly assessed for groups affected by siting decisions; they were assessed either for government decision-makers, or for representatives of interest groups.

Keeney and Nair (1974) and Fisher and Ahmed (1974) have discussed the use of utility theory for siting power plants, though without actually reporting application of the method. Dee et al. (1973) have developed an "environmental evaluation system" for water resource projects, which is a set of non-linear single-attribute utility functions over 78 attributes of environmental impact which are aggregated by a weighted sum, of the form of Equation 6, and thus in essence is a multi-attribute utility function for environmental impacts of the additive form.

5. *Advantages and Disadvantages of Preference Methods*

The advantages of utility analysis over the methods previously discussed spring from its rigorous handling of preference for impacts and uncertainty. It is the only one of the evaluation methods that adequately accounts for dependence among the desirabilities of different impacts and for uncertainty in impact predictions. The method allows differences in desirability as perceived by different groups to be introduced, and theory is currently being developed to incorporate varying group utility functions analytically in decision-making (Kirkwood, 1974).

The disadvantages derive mainly from problems of application: assessing utility, dealing with sometimes messy mathematics, and lack of conceptual simplicity. The problem of coalescing the utility functions of different groups into one function is more explicit with utility models, but is a problem inherent in siting and not in a particular method. Other methods either ignore this question or treat it judgementally. Perhaps the major problem is measurement: what are we measuring when we assess over large groups, and does whatever we measure accurately reflect individuals' "true" preferences or merely their monetary whims? The procedures of utility assessment seem better on this point than opinion survey generally, as they confront a subject with decisions involving trade-offs among impacts rather than simply asking opinion-type questions; however, the objection of economists that surveys and market behaviour represent qualitatively different things and that the latter may be more valid and reliable still plagues the effort. The answer to this problem is not immediately apparent, and certainly a closer look at the measurement problem might prove more helpful than much of the current effort to expand the mathematical base of utility theory.

APPENDIX A

Utility Assessment[1]

The assessment of utility functions involves having the subject whose preferences are to be assessed choose among various alternatives with uncertain and certain outcomes; then an interval scaling of his preferences is back-figured from his answers. As an example, consider the choice between a certainty of receiving \$5,000, and the wager with equal chances of winning \$10,000 and \$0. For convenience, we scale the utility function so that u(\$10,000) = 1 and u(\$0) = 0. The expected utility value of the wager is

$$0.5\ U(\$10{,}000) + 0.5\ u(0) = 0.5 \quad .$$

If the subject chooses the sure \$5,000 over the wager, then we can infer that the utility of \$5,000 must be greater than the expected utility of the wager, which is 0.5. Similarly if the subject, faced with the choice between \$3,000 and the wager, chooses the wager, then the utility of \$3,000 must be below 0.5. Questioning would continue until a value is established for which the subject is indifferent.

A similar procedure would be used in multi-attribute problems. A series of choices is presented to establish whether preference independence properties hold, and whether a sum or product form is appropriate. If either is appropriate, the problem reduces to assessing single-attribute scalings, followed by simple multi-attribute questions to obtain scaling constants among impacts. If the simple forms are not appropriate, more complicated series of questions must be used.

APPENDIX B

Axioms of Utility Theory

Utility function analysis depends on seven axioms. Before stating them, it is helpful to define some notation. A simple lottery, written $L(x_1,p,x_2)$, is the event where there is a chance p that x_1 will occur and a chance $1 - p$ that x_2 will occur. The symbol > means that, when faced with the choice between the event to the right and that to the left of the symbol, the latter is preferred. The symbol $\sim$ means that the decision-maker is indifferent to the choice between the two events, and $\lesssim$ means that the event to the left is not preferred to that on the right. Thus, the statement $x_1 \sim L(x_2,p,x_3)$ says that the decision-maker is indifferent to the choice between the x_1 for certain, and the lottery yielding either x_2 with probability p or x_3 with probability $1 - p$. We can now formally state the axioms, based on those used in Pratt, Raiffa and Schlaifer (1965).

Axiom 1: *Existence of Relative Preferences.* For every pair of values x_1 and x_2, the decision-maker will have preferences such that either $x_1 \sim x_2$, $x_1 > x_2$, or $x_2 > x_1$.

Axiom 2: *Transitivity.* For any lotteries L_1, L_2, and L_3, the following holds:

i) if $L_1 > L_2$ and $L_2 > L_3$ then $L_1 > L_3$

ii) if $L_1 \sim L_2$ and $L_2 \sim L_3$ then $L_1 \sim L_3$

and so on.

1. Full descriptions of utility assessment can be found in Schlaifer (1959). Practical assessments are discussed in Gros (1974) and Keeney (1972). Also, interactive computer programs are available (Schlaifer, 1971; Keeney & Sicherman, 1975).

Note that any deterministic value x_i can be expressed as a degenerate lottery, so Axiom 2 requires transitivity between deterministic events also.

Axiom 3: *Comparison of Simple Lotteries.* If for the decision-maker $x_1 > x_2$, then

i) $L_1(x_1,p_1,x_2) \sim L_2(x_1,p_2,x_2)$ if $p_1 = p_2$,

ii) $L_1(x_1,p_1,x_2) > L_2(x_1,p_2,x_2)$ if $p_1 > p_2$.

Axiom 4: *Quantification of Preferences.* For each possible consequence x, the decision-maker can specify a number $\pi(x)$, $0 \lesssim \pi(x) \lesssim 1$, such that $x \sim L(x^*,\pi(x),x_*)$, where x^* is the most preferred and x_* the least preferred outcome. The value $\pi(x)$, the indifference probability of the lottery, is a measure of utility.

Axiom 5: *Quantification of Judgemental Uncertainties.* For each possible event E which may affect the consequence of a decision, the decision maker can specify a probability P(E), $0 \lesssim P(E) \lesssim 1$, such that he is indifferent between $L(x^*,P(E),x_*)$ and the situation where he receives x^* if event E occurs and x_* if it does not.

Axiom 6: *Substitutability.* If a decision problem is modified by replacing one lottery or event by another which is equally preferred, then he should be indifferent between the old and the modified decision problems.

Axiom 7: *Equivalence of Conditional and Unconditional Preferences.* Let L_1 and L_2 designate lotteries that are possible only if event E occurs. After it is known whether or not E occurred, the decision-maker must have the same preference between L_1 and L_2 as he had before it was known whether E occurred.

VII *Conclusions*

We have reviewed three methodologies which apply multi-objective decision techniques to site selection problems for large constructed facilities. Our major observations are the following.

1. The methodologies are distinguished by having different objective functions. One must be aware of the assumptions underlying objective functions, and select that which best fits the decision problem considered.

2. Only certain mathematical operations on preference measures are permissible. One should keep in mind the scale on which preference measures have been made, and the mathematical operations that are appropriate. Failure in this respect can result in numbers that have no interrelational meaning.

3. Sensitivity analyses should always be performed. Uncertainty in the parameter values of the objective function, along with uncertainties in impact prediction, lead to uncertainties in objective function values. One should check how sensitive results are to these uncertainties.

4. Siting decisions are inherently political. The analyst's role in this process should be to eliminate all but the two or three "best" sites, and then to detail impacts for these, aggregated against the major objectives of cost, environmental degradation and social disruption.

References

1) Ackoff, R.L., S.K. Gupta and J.S. Minas. Scientific Method: Optimising Applied Research Decisions. Wiley, N.Y. 1962.

2) Baecher, G.B. "Sampling for Interest Group Preferences." IIASA RM-75-00. International Institute for Applied Systems Analysis, Laxenburg, Austria. Forthcoming 1975.

3) Beer, P. "Quantification of Siting Factors for Rational Selection of Alternative Sites." ANS Conference on Power Plant Siting, 1974, p. 12, ff.

4) Berkman, R.L. and W.K. Viscusi. Damming the West. Grossman Publications, N.Y. 1973.

5) Bishop, A.B. "Approach to Evaluating Environmental, Social and Economic Factors". Water Resources Bull. 8, 4 (August 1972), 724.

6) Broadway, Robin W. "The Welfare Foundations of Cost-Benefit Analysis." The Economic Journal (December 1974), 926-939.

7) Brewer, G.D. Politicians, Bureaucrats and the Consultant: A Critique of Urban Problem-Solving. Basic Books, N.Y. 1973.

8) Buehring, William A. "A Model of Environmental Impacts from Electrical Generation in Wisconsin." Ph.D. Thesis Title (tentative), University of Wisconsin, Nuclear Engineering Department, August 1975.

9) Burnham, J.B. et al. "A Technique for Environmental Decision Making Using Quantified Social and Aesthetic Values." BNWL-1787, Battelle Pacific Northwest Laboratories, Richland, WA. February 1974.

10) Collings, J.B. "Urban Priorities and Attitudes: Socio-Economic and Lifestyle Determinants." Vancouver Urban Future Project, Report 6, University of British Columbia, 1973.

11) Dee, N. et al. "An Environmental Evaluation System for Water Resrouces Planning." Water Resrouces Research, 9 (1973), 523-539.

12) deNeufville, R. and R.L. Keeney. "Use of Decision Analysis in Airport Development for Mexico City." In Analysis of Public Systems. Eds. A.W. Drake et al. MIT Press, Cambridge, MA. 1972.

13) Dorfman, R. "Formal (Mathematical) Models in the Design of Water Resources Systems." Water Resources Research, 1 (1965), 329.

14) Dorfman, R. Ed. Measuring Benefits of Government Investments. Brookings Institution, Washington, D.C. 1965, pp. 429.

15) Dorfman, R. "Basic Economic and Technologic Concept: A General Statement." In A. Maass et al. Design of Water Resource Systems. Harvard University Press, Cambridge, MA. 1962.

16) Eckstein, P. "A Survey of the Theory of Public Expenditure Criteria." In Ed. J.M. Buchanan, Public Finances: Needs, Sources and Utilisation. Princeton University Press, Princeton, N.J. 1961.

17) Eckstein, O. Water Resources Development. Harvard University Press, Cambridge, MA. 1958.

18) Edwards, W. and A. Tversky, Eds. Decision Making. Penguin, London, 1967.

19) Fishburn, P.C. Utility Theory for Decision Making. Wiley, N.Y. 1970.

20) Fishburn, P. Decision and Value Theory. Wiley, N.Y. 1964, p. 451.

21) Fisher, A.C. and J.V. Krutilla. "Valuing Long Run Ecological Consequences and Irreversibilities." Journ. Environmental Economics and Management, 1 (1974), 96-108.

22) Fischer, J.A. and R. Ahmed. A Systematic Approach to Evaluate Sites for Nuclear Power Plants. Transactions of the ANS Conference on Nuclear Power Plant Siting, August 25-28, 1974, Portland, OR. pp 7-8.

23) Goeller, B. "San Diego Clean Air Project." Rand Report R-1362-DS, Santa Monica CA. 1974.

24) Gros. Jacques. "A Paretian Environmental Approach to Power Plant Siting in New England." Ph.D. Thesis, Harvard University, May 1974.

25) Hafele, W. "Hypotheticality and the New Challenges: The Pathfinder Role of Nuclear Energy." Minerva, 12, 3 (1974), 303-322.

26) Hansen, M.H., W. Hurwitz and W. Madow. Sample Survey Methods and Theory. Wiley .N.Y. 1953.

27) Heath, J.B. "Cost-Benefit Analysis and Airport Location." In Ed. M.G. Kendall. Cost-Benefit Analysis. The English Universities Press, London, 1971, pp. 395-304.

28) Hill, Morris. Planning for Multiple Objectives. Monograph / 5, Regional Sci. Res. Inst., Philadelphia, PA. 1973.

29) Holling, C.S. "Resilience and Stability of Ecological Systems." IIASA RR-73-3. International Institute for Applied Systems Analysis, Laxenburg, Austria, 1973.

30) Hyman, H.H. Interviewing in Social Research. University of Chicago Press, Chicago, IL. 1954.

31) Jones, G.R., B.A. Gray and J.B. Buniham. A Method for the Quantification of Aesthetic Values for Environmental Decision Making. Transactions of the ANS Conference on Nuclear Power Plant Siting, August 25-28, 1974, Portland, OR., pp. 49-50.

32) Joskow, P.L. "Approving Nuclear Power Plants: Scientific Decision of Administrative Charade?" Bell Journal of Ec. & M. Sc., 5 (1974), 320-332.

33) Keeney, R.L. "Concepts of Independence in Multiattribute Utility Theory." In Eds. J.L. Cochrane and M. Zeleny. Multiple Criteria Decision Making, 1973, pp. 62-71.

34) Keeney, R.L. "Multiplicative Utility Functions." Technical Report No. 70, Operations Research Centre, MIT, Cambridge, MA., March 1972.

35) Keeney, R.L. and K. Nair. "Decision Analysis for the Siting of Nuclear Power Plants: The Relevance of Multiattribute Utility Theory." Technical Report No. 96, Operations Research Center, MIT, Cambridge, MA., June 1974.

36) Keeney, R.L. and H. Raiffa. Decision Analysis with Multiple Conflicting Objectives. Wiley, N.Y. (forthcoming).

37) Keeney, R.L. and A. Sicherman. "An Interactive Computer Program for Assessing and Analysing Preferences Concerning Multiple Objectives." IIASA RM-75-12, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1975.

38) Kendall, M.G. Cost-Benefit Analysis. The English University Press, London, 1971.

39) Kirwood, C.W. Group Decision Making Under Forced Uncertainty. Technical Report No. 74-5, The Department of Industrial and Operations Engineering, The University of Michigan, Ann Arbor, Michigan, October 1974.

40) Koopmans, T.C. "Stationary Ordinal Utility and Impatience." Econometrica, 23 (1960), 287-309.

41) Krutilla, J.V. Ed. Natural Environments: Studies in Theoretical and Applied Analysis. Johns Hopkins Press, Baltimore, Md., 1972.

42) Krutilla, J.V., C.J. Cicchetti, A.M. Freemen III and C.S. Russell. "Observations on the Economics of Irreplaceable Assets." In Eds. A.V. Kneese and B.T. Bower. Environmental Quality Analysis. Johns Hopkins Press, Baltimore, Md., 1972.

43) Layard, R. Cost-Benefit Analysis: Selected Readings. Penguin, 1972.

44) Lee, D.B. Jr. "Requiem for Large-Scale Models." Journal of the American Institute of Planners, 39, 3 (1973), 163-78.

45) Leopold, L.B., F.E. Clarke, B.B. Hanshaw and J.R. Balsley. "A Procedure for Evaluating Environmental Impact." Geological Survey, Circular No. 645 (1971), 1-13.

46) Lichfield, N. Cost-Benefit Analysis in Planning: A Critique of the Roskill Commission. 'Regional Studies 5 (1971), 157-183.

47) Lichfield, N. "Evaluation Methodology of Urban and Regional Plans: A Review." Regional Studies, 4 (1970), 151-165.

48) Lichfield, N. "Economics in Town Planning." Town Planning Review, 39 (1968), 13-15.

49) Maass, A. et al. Design of Water-Resource Systems. Harvard University Press, Cambridge, MA., 1962.

50) MacCrimmon, K.R. Decision Making Among Multi-Attribute Alternatives: A Survey and Consolidated Approach. RAND RM-5877 DOT, Santa Monica, CA. 1968.

51) MacCrimmon, K.R. and M. Toda. The Experimental Determination of Indifference Curves. The Review of Economic Studies, 36 (1969), 433-450.

52) McKean, R. Efficiency in Government Through SA with Emphasis on Water Research Development. Wiley, N.Y., 1958.

53) Majone, G. On the Logic of Standard Setting in Health and Related Fields. Paper presented at the International Institute for Applied Systems Analysis conference on "Systems Aspects of Health Planning" held at Baden, Austria, August 20-22, 1974.

54) Major, D.C. Multiobjective Redesign of the Big Walnut Project. Systems Planning and Design, Prentice-Hall, Englewood Cliffs, N.J. 1974.

55) Manheim, M.L. and F. Hall. Abstract Representation of Goals. Department of Civil Engineering, MIT, Cambridge, MA., 1967.

56) Marglin, S.A. Public Investment Criteria. MIT Press, Cambridge, MA., 1967.

57) Marglin, S.A. "Objectives of Water-Resource Development: A General Statement." in A. Maass et al. Design of Water-Resource Systems, Harvard University Press, Cambridge, MA., 1962.

58) Meyer, R.F. "On the Relationship Among Utility of Assets, The Utility of Consumption and Investment Strategy in an Uncertain, but Time Invariant World." Proc. 4th IFORS Conference, Venice, Italy, 1969.

59) Mishan, E.J. Cost-Benefit Analysis. Geo. Allen & Unwin Ltd., London, 1971.

60) Olson, M. The Logic of Collective Action. Harvard University Press, Cambridge, MA., 1965.

61) Otway, H.J., R.K. Lohring and M.E. Battat. "A Risk Estimate for an Urban-Sited Reactor." Nuclear Technology, 12 (1971), 173-184.

62) Pratt, J.W., H. Raiffa and R.O. Schlaifer. Introduction to Statistical Decision Theory. McGraw Hill, New York, N.Y., 1965.

63) Prest, A.R. and R. Turvey. "Cost-Benefit Analysis: A Survey." Economic Journal (1965).

64) Raiffa, H. Decision Analysis. Addison-Wesley, Reading, MA., 1968.

65) "Roskill Commission on the Third London Airport." Her Majesty's Stationery Office, London, 1970.

66) Ross, M. "London's Third Airport: Quantitative Decision-Making." Ekistics, 218, (1974), 27-44.

67) Schlaifer, R.O. Computer Programs for Elementary Decision Analysis. Harvard Cambridge, MA., 1971.

68) Schlaifer, R.O. Probability and Statistics for Business Decisions. McGraw-Hill, New York, N.Y., 1959.

69) Sewell, W.R.D. Broadening the Approach to Evaluation in Resources Management Decision-Making. Journal of Environmental Management, 1 (1973), 33-60.

70) Smith, N.M. Jr. "A Calculus for Ethics: A Theory of the Structure of Value." Behavioural Science, 1 (1956), 111-186.

71) Stanley, J.K. "A Cardinal Utility Approach for Project Evaluation." Socio-Econ. Plan. Sci., 8 (1974), 329-338.

72) Starr, C. "Benefit-Cost Relationships in Socio-Technical Systems." Paper presented at the Symposium on Environmental Aspects of Nuclear Power Stations, UN Headquarters, New York, N.Y., August 14, 1970.

73) Stevens, S.S. "Measurement, Psychophysics, and Utility." In Eds. C.W. Churchman and P. Ratoosh. Measurement: Definitions and Theories, Wiley, New York, N.Y., 1959.

74) Stevens, S.S. and E.J. Galanter. "Ratio Scales and Category Scales for a Dozen Perceptual Continua." Journal of Experimental Psychology, 54 (1958), 377-411.

75) Swain, H. "Evaluating Growth Proposals." IIASA internal paper, April 1975.

76) Ting, H.M. "Aggregation of Attributes for Multiattribute Utility Assessment." MIT Center, Technical Report /66, MIT, Cambridge, MA., 1971.

77) UNIDO. "Guidelines for Project Evaluation." United Nations, New York, N.Y. 1972.

78) US Atomic Energy Commission. "Directorate of Regulatory Standards." Preparation of Environmental Reports for Nuclar Power Plants, Regulatory Guide 4.2, Washington, D.C., March 1973.

79) Walterns, C.J. "Foreclosure of Options in Sequential Resource Development Decisions." IIASA RR-75-12, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1975.

80) US Water Resources Council. "Proposed Principles and Standards for Planning Water and Related Land Resources." US Federal Register, 36, 245, Part II (1971).

81) US Water Resources Council. "Procedures for Evaluation of Water and Related Land Resource Projects." US Government Printing Office, 1965.

82) Webb, E.J., D.T. Campbell, R.D. Schwartz and L. Sechrest. Inobtrusive Measures, Rand McNally, Chicago, Il. 1972.

83) Weisbrod, B.A. "Income Redistribution Effects and Benefit-Cost Analysis." In Ed. S.B. Chase, Jr. Problems in Public Expenditure Analysis. The Brookings Institute, 1968.

DEREK W. BUNN
HOWARD THOMAS

PROBLEMS IN THE APPLICATION OF DECISION ANALYTIC METHODS TO POLICY FORMULATION

This paper provides a review of the major problems faced in adapting the decision analysis paradigm to the policy situation. Currently, it is felt that with reasonable adaptions the state of the art in decision analysis is sufficient to provide meaningful analytic tools for policy-makers.

4. Problems in the Application of Decision Analytic Methods to Policy Formulation by Derek W. Bunn and Howard Thomas

The purpose of this second part is to examine more closely some of the specific practical issues which are usually encountered in the application of decision-analytic methods to policy formulation. The theory of decision analysis provides rational principles for the ideal (i.e. "coherent") individual faced with a nicely structured problem involving well-understood outcomes and uncertainties. In the wider context of organisational policy formulation, however, the problem is not so well-defined. Decision theory still provides the rational paradigm when the organisation is considered as a single entity, but several complicating aspects have now to be dealt with in a formal way.

There is the problem of multiple conflicting objectives which Baecher, Gros and McCusker reviewed in Part 1, and this is a constant theme through most of the papers in this volume.

The optimum policy should also represent a consensus both amongst the set of decision-makers (the "expert resolution problem") and the set of groups affected differently by the policy (the "multiple impact group problem"). Williams in Part 1 emphasised the importance of overcoming these problems if the decision analysis method is to adequately formalise policy-making.

The uncertainties have to be dealt with in a more structured way than by just assessing individual subjective probability distributions. A formal synthesis of all the forecasts, opinions and other indications has to be attempted (c.f. Bunn (6)).

The problem must be structured in such a was as to truly reflect any time-sequenced dependencies and spatial heterogeneity in the outcomes. The decision-tree method will often be inadequate to elucidate the complexity and number of options under consideration. In such cases, the optimisation methods of mathematical programming may have to be introduced in order to identify the optimum.

Hence, a thorough decision analysis of even a relatively small set of options is frequently extremely complex and time consuming. It is important in practice, therefore, to undertake the maximum amount of preliminary screening to reduce the number of options under consideration to a minimum.

Some of the key issues are therefore:

- Screening
- Probability Assessment
- Consensus and Expert Resolution
- Multiple Conflicting Objectives
- Structuring and Optimisation

These will now be considered in more detail.

Screening

Screening procedures could be distinguished into those which attempt to reduce the set of options and those which attempt to simplify the structure of the decision model. Conceputally it might appear that these are two distinct stages in a rational decision analysis. The options are first reduced to a minimum and the final decisión model is then simplified to the most realistic structure. However, options cannot adequately be 'screened' without a simple decision model and furthermore, in 'screening' the structure of that model, extraneous options would fall out anyway as additional complications. This simultaneity in screening options and structural assumptions is recognised in the paper of Byer and de Neufville which is included in this Part.

A fairly common type of multiattribute screening model is a simple extension of the Achievement Matrix approaches which are quite thoroughly reviewed in Baecher, Gros and McCusker. In the goals-achievement matrix of, for example, Hill (14), the matrix consists of a set of subjectivly assessed scores for each option (defining the rows) against each relevant dimension (defining the columns). On assigning relative weights to each dimension, the weighted score can then be computed for each option. In formal terms, if s_{ij} denoted the score of option i on the j^{th} attribute, and w_j the relative weight given to attribute j, then the score S_i, where

$$S_i = \sum_{\forall j} s_{ij} w_j$$

is used to rank the options. This is also often known as the Churchman-Ackoff procedure and has been widely used, particularly in evaluating research and development projects, e.g. Williams (34), Thomas (30).

The main value of this Churchman-Ackoff procedure is in its intuitive appeal as a simple formalisation of a decision situation, and this probably accounts for its popularity in practice. As a decision model, however, it does assume that the attributes are considered independent and that preferences are adequately represented by the implicitly linear scoring measure. It is often argued, however, that even if these assumptions are too strong to provide a valid ranking of all the options, they may still be sufficiently robust for screening out the top 'subset' so that more detailed analysis and synthesis may be carried out on them. The value of the approach put forward by Byer and de Neufville lies in the very fact that these such assumptions are explicitly tested and more importantly that all this is done whilst taking into account uncertainty in the outcomes. Another drawback with the Churchman-Ackoff procedure, is, in fact, its lack of attention to uncertainty and risk. Evidently "risk" could be treated as one of the attributes but that would beg the question of its definition. Even if its measurement could be reduced to a single value, such a value would be needed for each of the uncertain attributes. Thus even a univariate measure of risk may double the number of columns in the matrix.

Screening for risky ventures has received considerable attention in the literature of capital-budgeting and financial investment analysis, but in all cases on a single finanacial outcome attribute. Screening under uncertainty has apparently not been extended to the multiattribute domain and incorporated in the Achievement-Matrix approaches.

The simplest option screening under uncertainty equates risk with the variance of the outcome, and the issue is seem as one of balancing mean against variance. Screening consists of identifying the "efficient set" of options. The efficient set excludes all "dominated" options and an option is said to be dominated in the set if there exists either another option with the same mean but lower variance, or with the same variance but higher mean. This approach has been most evident in the literature on Portfolio Theory, for example Markowitz, (16).

Risk is sometimes interpreted as the probability of a "disaster" or risk of "ruin" and often thereby assessed as the probability of the outcome being less than a certain critical value (see for example, Broyles and Thomas (3)). In the investment context where the outcome dimension is a Net Present Value measure, this critical value is often taken as zero. If the uncertainty in each option is furthermore described by a Normal probability distribution function, then the option with minimum risk is that with the highest mean/standard derivation ratio.

The mean-variance approach does in fact assume that either all the options have Normally distributed outcomes, or that the decision-maker's utility function on the outcomes is of quadratic form. For screening purposes, it is a question of how robust these assumptions are in identifying the best subset.

Less restrictive constraints on the form of implicit function can be dealt with using screening models developed from the concept of "stochastic dominance". Stochastic dominance is said to occur if the expected utility of an option is greater than that of another over a whole class of utility functions. The theory owes much of its development to Hadar and Russell (11), Hanoch and Levy (13) and Whitmore (33). A recent survey article is that by Eilon and Tilley (9). The set of conditions derived from the stochastic dominance concept is given below.

First Degree Stochastic Dominance (FSD)

This makes very weak assumptions on the form of the utility function: only that $U(x)$ is finite, continuously differentiable, and strictly increasing over x. If $F_1(x)$, $F_2(x)$ are the distribution functions for two options, then if:

$$F_1(x) \leq F_2(x)$$

option 1 dominates option 2 (except, of course, in the case of equality over all x) in the sense of FSD.

Second Degree Stochastic Dominance (SSD)

This assumes a risk averse (concave) utility function over x, in addition to the assumption of FSD. If

$$\int_{-\infty}^{z} (F_2(x) - F_1(x))\, dx \geq 0 \qquad \forall z$$

then option 1 dominates option 2 in the sense of SSD. This is more restrictive than FSD and can therefore, when appropriate, further screen an efficient set derived from FSD.

Third Degree Stochastic Dominance (TSD)

This test requires in addition to FSD, that $U'(x) > 0$, $U''(x) > 0$ and $U'''(x) > 0$ for all x. If

$$\int_{-\infty}^{y} \int_{-\infty}^{z} (F_2(c) - F_1(x))\, dxdz \geq 0 \qquad \forall y$$

then option 1 dominates in the sense of TSD. This can further screen an efficient set derived from SSD, but the assumptions of TSD do not have an intuitively obvious appeal.

Indeed, it is a prerequisite of a screening model that it should be a considerably simpler model than the thorough Expected Utility approach which is in any case envisaged as being used on the final screened set. Furthermore, the procedure should ideally have great intuitive appeal since one of the more important benefits of screening before an expected utility analysis is in the sensitisation afforded on the conflicting issues involved in the problem. The decision-maker can gain a wider appreciation of the problem by working through a simple screening procedure and is then better prepared and more highly motivated to tackle the assessment tasks involved in deriving the more thorough utility and probability measures than if he had approached them 'cold'.

An extension of this reasoning suggests that if assessments from a panel are being used, it is better for them to work through the screening procedures individually and only later undertaking the final analysis as a group. The argument for working through a full expected utility analysis as a group entity is that it is easier to discuss and reconcile differences of opinion about the more meaningful and sensitive inputs into a decision model (forecasts, attitudes to risk, etc) than on the output criterion which in itself does not often have much tangible meaning to them (e.g. Expected Utility).

An ideal screening procedure would therefore represent a balance on being relatively simple and quick, sufficiently robust and intuitively appealing. The importance of the procedure as a preliminary familiarisation device before an Expected Utility analysis, apart from its basic function of simplifying the problem, should not be underestimated.

Finally it is worth mentioning that the development of sensible screening procedures can be seen as more than just aimed at facilitating the practical aspects of decision analysis. Baecher, Gros and McCusker in Chapter 2, for example, argue that since major policy decisions are essentially political, the role of decision analysis should be to present a final subset of good options for political evaluation. If this is the case, the whole decision analysis is effectively directed towards a screening procedure.

Probability Assessment

This section will be rather more detailed as this particular topic is one which is not specifically covered elsewhere in the volume. Useful survey articles on this topic are given by Hampton, Moore and Thomas (12) and Moore and Thomas (17). The line of analysis here will be firstly to discuss methods for estimating the subjective probability of a single realisable proposiation such as "it will rain tomorrow". This will then be extended to the estimation of a distribution over a countably infinite number of propositions. Attention will then be focussed on the assessment of certain conjugate distributions which are necessary in the analysis of important data-generating processes. Finally, a discussion of wider issues of predictive bias and some behavioural factors of implementation will follow.

The Estimation of the Probability of a Realisable Proposition

Since the probability of a proposition will rarely be estimated without the implicit consideration of its complement, this section will more strictly refer to the assessment of sets of realisable propositions. However, discussion will be restricted to sets of propositions of a size applicable to event 'fans' on a decision tree, but not large enough to be evaluated as "continuous" distributions.

The apparently simplest assessment procedure would be one in which the subject responded by directly articulating a set of numerical probabilities. Unfortunately psychologists (c.f. Phillips (22)) have suggested that this is not necessarily the simplest conceptual task, particularly for a decision maker with little experience in probability. For example, different measures are obtained if the individual responds in terms of odds ratios or direct probabilities. For this reason, the use of standard devices is generally advocated as a medium of expression, some of which are described in Bunn and Thomas (7).

It will be useful to conceptualise a subject's fundamental notions on uncertainty to be in the form of a Non-Probabilistic Chance Perception (NPCP) and that the function of the assessment procedure is one of mapping this cognitive structure into a consistent Probability Density Function (PDF). The imposition of the consistency requirement may also introduce a degree of belief formulation during the assessment procedure itself, apart from its pure operation as a transformation. Thus, standard devices attempt to furnish the individual with a physical equivalent to their NPCP from which a PDF can then easily be deduced.

A useful standard device is an urn filled with 1000 identically shaped balls. Each ball is identified with a number, from 1 to 1000. The simple experiment of drawing, blind, one ball from the urn is to be performed. Phillips and Thomas (25) describe this method as it is often presented by a decision analyst in practice:

> To see how the standard device can be used to measure degrees of belief, we must consider two bets, one involving the event whose probability you wish to assess, and one involving the

standard device.

Suppose, for example, you want to determine the probability that it will rain tomorrow. Imagine that the following bet has been offered to you:

BET A < If it rains tomorrow, you win £1,000
If no rain tomorrow, you win nothing.

The tree-diagram of figure 1.A is a convenient representation of this bet.

Figure 1 Tree-Diagram for the 'rain-tomorrow' bet and for the reference bet.

Now imagine that balls 1 through 500 in the standard urn have been painted red while the remaining 500 balls have been coloured blue. The balls are thoroughly mixed, and one is to be drawn by a blindfolded observer as tomorrow draws to a close. Now consider this bet:

BET B < If the ball drawn is red you win £1,000
If the ball is blue, you win nothing.

This bet is shown in Figure 1.B. We would all agree that the probabilities of drawing a red or blue ball are 0.5 respectively, and these probabilities are shown on the branches of the tree. Remember, we are trying to find out what probabilities should be shown on the branches of the tree representing the 'rain tomorrow'.

Consider both bets. Which do you prefer, A or B? Suppose you prefer B. Then there must be a better chance for you to win £1,000 with Bet B than with Bet A. Thus, the probability of rain tomorrow in your judgement, is clearly less than 0.5.

By changing the proportion of red balls in the urn, it is eventually possible to find a mix of red and blue balls that make you indifferent between the two bets. When this point is reached, then we are justified in assigning the same probability to the event 'red ball is drawn' as we are to the event 'rain tomorrow'. At no time is it necessary to ask a question more complex than 'Do you prefer this bet or that one, or are you indifferent between them?' Numerical measurement of an individual's subjective probability can thus be obtained simply by asking questions of preference.

Other standard devices have been popular. A pie diagram, or spinner, is a favourite with Stanford Research Institute. A circle is divided into two sectors and the relative sizes of the sectors can be adjusted. A spinner randomly selects one of the two sectors and hence the larger a sector the greater its chance of being chosen. The same bets as those shown in Figure 1 can be offered, but the outcomes for Bet B are determined not by drawing a ball from an urn, but by noting which sector is chosen. The relative sizes of the sectors are adjusted until the indifference point is reached; the sector sizes then represent the probabilities of the event being assessed and its complement.

The pre-requisite of a standard device is that it should have easily perceived probabilistic implications, otherwise it will introduce bias. Thus Phillips and Thomas (25) report some preliminary investigations which suggest that assessments using the urn device are 0.02 to 0.07 larger than the probabilities from the SRI spinner.

It is proposed to distinguish between task bias which is characteristic of the assessment method itself and conceptual bias which is idiosyncratic to the individual. Task bias could be caused by the standard device having misunderstood probabilistic implication or because it also structures thinking and maybe changes the fundamental beliefs of the individual in some systematic way. The more fundamental conceptual bias represents a faulty NPCP and will relate to his inability to process information and deduce the causal implications of the various inductive hypotheses. The endeavour to develop formal methods for synthesising inductive models would be an attempt to reduce this conceptual bias, but fundamental limitations relating to an individual's perception of chance processes militate against this and the reduction of task bias.

Tversky and Kahnemann (32) have recently presented an important paper dealing with a characterisation of different sorts of conceptual bias. They isolate three types of systematic bias in the formulation of probabilistic judgement; representativeness, availability and adjustment.

Representativeness

Individuals apparently formulate probabilistic judgement by means of a representativeness heuristic. Thus if x is considered highly representative of a set A, then it is given a high probability of belonging to A. However, this approach to the judgement of a likelihood leads to serious bias because many of the factors important in the assessment of likelihood play no role in such judgements of similarity. One factor is the prior probability or base rate frequency. For example, given a neutral description of a person and being asked to estimate the probability of him being a lawyer or an engineer, subjects were found to answer 0.5 regardless of prior information on the relative numbers of lawyers and engineers in the population. Similarly, the representativeness heuristic does not take any account of sample size. Thus the manifestation of the gamblers fallacy can be ascribed to the belief that randomness is expected to be represented in even very small samples. Tversky and Kahnemann describe many fascinating cases of such bias.

Availability

Reliance on the availability heuristic introduces bias through the inadequacy of the cognitive process in conceptualising all of the relevant information. There is a memory retrievability problem which can cause a bias such as the probability of a road accident increasing dramatically after witnessing such an event and by much more than by just reading about it. The limitations of the memory search process cause people to judge that there are more words beginning with 'r' and with 'r' in the third place when in fact the converse is true. Conceptual limitations of imaginability and scenario formulation encourage subjects to believe for example that many more committees can be constructed of size 2 from 10 than of size 8. Again Tversky and Kahnemann relate many other interesting examples.

Adjustment and Anchoring

In most situations it is found that individuals formulate their general belief structure by starting from some obvious reference point and adjusting for special features. Typically, however, the adjustment is not sufficient and a bias towards these initial values is described as anchoring. Thus when subjects were asked to estimate within

5 seconds the product 8x7x6x....x1 they gave a much higher answer than those asked the product 1x2x3....x8. A much fuller consideration of anchoring will follow in the context of the fractile assessment method for a distribution in the next section.

There is now an enormous amount of published experimental work investigating how individuals deviate from Bayesian rationality in formulating their probabilistic judgement. Much of it is equivocal and the paper by Tversky & Kahnemann (32) represents one of the few attempts at deriving an overall theory. Ward Edwards and his followers (c.f. Phillips and Edwards (23)) looked not so much at the formulation of belief but in its revision. Compared with the rational Bayesian paradigm, individuals have generally been found to be "conservative" information processors, underestimating the overall diagnosticity of observed evidence. Again, more will be said about conservatism bias in the next section in the particular context of the 'imaginary results' assessment method.

One way of minimising some of this bias is to ask for less precise estimates by not requiring the responses to be on a metric scale. There have been various forms of psychometric ranking methods proposed. In some cases a simple ranking of the outcomes may be sufficient or a decision analysis can be structured in such a way as to require only a sensitivity analysis of certain crucial probability assessments. However, in general, the necessary degree of precision will imply a ranking of first differences, as in Smith (28).

The indirect estimation of probabilities from gambling preferences makes strong behavioural assumptions. Quite often it is assumed that the individual is behaving to maximise expected monetary value. Thus, returning to the gamble presented in Figure 1A, a statement of his certainty equivalent (CE) for the gamble allows the probability to be imputed as CE/1000.

But to have any confidence in this as a predictive probability, it should be ascertained that his utility function is in fact linear over this range. However, if trouble is to be taken in measuring the individual's utility curve in the first place, there is no reason why the payoffs should not be appropriately mapped into utility in order to obtain consistent subjective probabilities. If a von Neumann & Morgenstern utility function is derived using standard devices to articulate the probabilities presented in the artifical gambles, then the subsequent use of this function in the derivation of subjective probabiltiies will give valid estimates providing the individual obeys the coherence axiom. Unfortunately, this may not be the case. Phillips (22) quotes experimental evidence from Slovic that individuals react differently in gambling situations. Some people pay more attention to the chance of winning, others to the chance of losing, while a further group seem to look mainly at the amounts of the payoffs.

Estimation of the Subjective Probability Distribution

The problem of assessing the distribution function over a countably infinite set of propositions is usually reduced to a set of discrete assessments requiring only the application then of one of the methods of the previous section. Bias can however be introduced according to the way in which the range of propositions is split up. For the usual fractile method of assessment, Morrison (18) designed the following questionnaire:

Qu.1. At what value of the variable, F(50), do you feel that there is a 50 per cent chance that the true value of the variable will be below F(50)? - thus establishing the value at which CDF = 0.5.

Qu.2. Given that the true value of the variable is below F(50) at what value of the variable F(25) do you feel there is a 50 per cent chance that the true value of the variable will be below this value? - thus establishing the value at which CDF = 0.25.

Qu.3. Given that the true value is above F(50) at what value of the variable F(75) do you feel there is a 50 per cent chance that the value of the variable will be below this value? - thus establishing the value at which CDF = 0.75.

Evidently, this method of successive medial bisection will result in the set of quartiles or octiles, etc. A more recently favoured fractile method (because of its potential tendency to reduce the anchoring effect) is to assess the tertiles, i.e. those fractiles which split the range into three equally probable intervals. The successive extension of this to yield noniles may be attractive and there is doubtless the prospect of a hybrid fractile method generating sextiles gaining certain appeal.

This family of fractile assessment methods appears the most convenient way to estimate the distribution function over a continuous range. It is to be preferred to the the direct estimation of a probability histogram over a set of pre-specified intervals on the range, which would be the basis for interpolating the PDF, since it does not involve a response in the form of a probability metric. In all the above fractile methods, the responses are in the form of equiprobable intervals.

Barclay and Peterson (2) compared the tertile method with the PDF histogram approach and found that anchoring bias was considerable more serious in the PDF. A central interval in the PDF method only captured the "true" value 39% of the time compared with the 75% of perfect calibration. In the tertile method, instead of the ideal 33 1/3%, the central interval captured the "true" value only 23%.

This tertile adjustment bias compared with the 33% for the 50% central interval in the earlier quartile experiments of Alpert & Raiffa (1). Tversky (31), Winkler (35) and Pickhardt & Wallace (26) have all reported similar anchoring bias in the quartile method.

However, it should be recognised that all this evidence is based upon artificial laboratory experimentation where the subjects will not have the same degree of motivation and personal involvement in the consequences of their probability estimates as they would in a real decision-making situation. There is an obvious need for more research in this area to be undertaken in terms of the 'real' decision-making processes of the individual where the possibility of other organisational or political biases may affect the conclusions drawn currently about the effectiveness of probability assessment procedures.

Stael von Holstein (29) was able to use professional investment analysts in their ongoing stock market forecasting and portfolio selection activity. He reported significant anchoring bias in the excessive tightness of the assessment distribution, although his choice of the PDF method may well have exacerbated this tendency.

Winkler and Murphy (32) were able to compare the quartile and PDF methods in the real world situation of weather forecasting. They reported that the PDF method exhibited greater anchoring bias than the quartile method for which in fact the central 50% interval captured the true value 47% of the time. They were fortunate however in having subjects with considerable experience in probabilistic forecasting; training and practice appears to have a very pronounced effect in reducing anchoring bias. Alpert and Raiffa (1), for example, found that after only one round in their experiment, the central 50% interval capture rate increased from 33% to 43% using the quartile method.

One factor which may reinforce anchoring bias is the importance generally placed upon self-consistency within the decision-maker's set of assessments. In a straightforward assessment method, it is easy for the subject to be pseudo-consistent precisely because he can perceive what he should believe in order to be consistent with his previous responses. In this way, his responses become firmly anchored from the starting point.

Some preliminary results have been published by Bunn (5) on a procedure which attempts to expose anchoring bias in the fractile method and thereby partially eliminate it. The method involves the derivation of an adjustment hysteresis effect where the responses are structured in such a way that it becomes difficult for the subject to exhibit pseudo-consistency. The results obtained so far are quite encouraging.

ESTIMATION OF SOME COMMON PRIOR DENSITY FUNCTIONS

Univariate distributions will first be considered. It should be emphasised that this class of density functions can always be parameterised by means of deriving the individual's subjective distribution as outlined previously and then fitting the required function to it. When an interactive computer package is available, particularly if it incorporates visual display, this may well be the best available procedure. Otherwise the problem is one of estimating the parameters in their own right.

The Normal PDF

Because of symmetry, the mean can be estimated as either the mode or median and, utilising the standard tables, the variance can easily be derived from any given set of fractiles.

The Beta PDF

The beta distribution is usually used to express prior opinion on the probability of one of the two dichotomous events in a Bernoulli process. If it is parameterised as

$$P(k) = B^{-1}(\rho+1,\nu+1)\ k^{\rho}(1-k)^{\nu-\rho}$$

where B(x,y) is the usual beta function and $\rho > \nu-1$, $\nu \geq \rho$, then the mode is given by ρ/ν and ν is equivalent to the number of previous trials which would be required to give the same precision on an empirical basis.

Thus the parameter ν can be assessed as an Equivalent Prior Sample (EPS) which the individual feels would be the empirical equivalent of his subjective opinion. Good (10) refers to this type of approach as one of imaginary results.

Another possibility is to elucidate how the individual's estimate of the mode would change on the basis of one more realisation. This is referred to as the method of Hypothetical Future Samples (HFS). If his estimate of the prior modal probability of 'success' is m* and the posterior m** is assessed after one further 'failure' is envisaged then

$$m^* = \nu/\rho$$

$$m^{**} = \nu/(\rho + 1)$$

which gives $\nu = m^{**}/(m^* - m^{**})$

Bayesian expectations, e* & e**, can be used instead of the modal point estimates if preferred. In order that the formulae would be directly analogous, it is suggested that the more usual parameterisation of the Beta distribution should be used in this case.

$$P(k) = B^{-1}(\rho,\nu)\ k^{\rho-1}(1-k)^{\nu-\rho-1}$$

with $\rho > 0$ and $\nu \geq \rho$.

Clearly, the methods of HFS and EPS make strong Bayesian assumptions about the way in which individuals process information. It was indicated earlier that individuals do in fact tend to be conservative processors of information. This tendency manifests itself in the assessment of too large a hypothetical sample, thus implying an excessively tight distribution similar to the anchoring bias in fractile assessment. Winkler (35) observed this in his experiments, although the subjects in his groups did report an intuitive preference for these imaginary results methods over the fractile and PDF methods.

Like most of the research in this area, our knowledge of conservatism bias is restricted to experimental behaviour in the laboratory. It is quite possible that this sort of bias could be largely situational and reflect the subject's unfamiliarity with the type of data generating processes and inferential tasks with which he is confronted. For optimal behaviour in these tasks, the subject may be very adept in dealing generally with the stationary Bernoulli process. The real world is characterised by non-stationarity and Phillips, Hays & Edwards (28) have remarked that the conservatism revealed in their experiments could be caused by the subjects believing that the data-generating process were non-stationary. Furthermore du Charme & Peterson (8) noted an improvement in the optimality of subjects when they were dealing with Normally generated data, which they suggested was due to their greater familiarity of this type of data from the real world.

Most of the work on cascaded inference has not succeeded in revealing significant conservatism bias. Models of cascaded inference attempt to formalise the more complex inferential tasks of the real world and in fact most of probability assessments were more optimal than in the simpler experiments where conservatism bias has been most evident. Bias, if anything, has tended to be excessive rather than conservative in this general context (c.f. Youssef and Peterson (38)).

THE INVERSE GAMMA DISTRIBUTION

This distribution is useful in the analysis of the Normal process, being the natural conjugate for the variance. It can be parameterised

$$f_{i\gamma}(y|\psi,\nu) = \frac{\exp(-\frac{1}{2}\nu\psi/y)\,(\frac{1}{2}\nu\psi/y)^{\frac{1}{2}\nu+1}}{\frac{1}{2}\nu\psi\Gamma\frac{1}{2}\nu}$$

for $y > 0$

The parameters can be given similar interpretations to those of the beta distribution. ν likewise represents the size of the hypothetical prior sample and the mode is given as

$$\nu\psi/(\nu+2)$$

Apart from using methods of imaginary results, a fractile method is possible. There is a standard result (c.f. Lavalle (15)) connecting the fractiles of a gamma distribution to the tabulated chi-squared distribution with the same degrees of freedom. If σ_p^2 denotes the p% fractile of the assessed variance distribution, then

$$\sigma_p^2/\sigma_q^2 = \chi^2_{(1-q)}/\chi^2_{(1-p)}\ \Big|_\nu$$

e.g.

$$\sigma^2_{.75}/\sigma^2_{.5} = \chi^2_{.5}/\ \chi^2_{.25}\ \Big|_\nu$$

and thus can be derived from the 'chi-squared' tables.

There are simple relations (c.f. Lavalle (15)) between the parameters of the inverse gamma (the natural conjugate for the variance of the normal process), the inverse gamma-2 (the natural conjugate for the corresponding standard deviation) and the gamma itself which is the natural conjugate for the parameter in the Poisson process.

Raiffa and Schlaifer (27) suggest that it is probably most convenient to assess the uncertainty in terms of the standard deviation and then translate it into variance or precision.

The Dirichlet Distribution

This is the natural conjugate to the multinomial process and can be expressed

$$f(\underset{\sim}{p}/\underset{\sim}{\rho},\nu) = \prod^{k} (p_i^{\rho_i - 1}) \prod^{k} (\Gamma \rho_i)/\Gamma\nu$$

with $\underset{\sim}{p}$ and $\underset{\sim}{\rho}$ k-dimensional vectors such that $\Sigma^{k} p_i = 1$, $\underset{\sim}{p} > 0$, $\Sigma^{k} \rho_i = \nu$ & $\underset{\sim}{\rho} > 0$.

The easiest conceptual approach is probably one of hypothetical results. ν represents the hypothetical sample size and ρ_i/ν the expected proportions for each element. ν can also be estimated according to the HFS method described for the beta distribution.

The fact that each element is marginally distributed beta means that a fractile assessment method is applicable. Each element can be marginally assessed as a beta subject to the constraint that ν should be the same for each and the ρ_i sum to ν.

The Inverted Wishart Distribution

The inverted Wishart distribution is the natural conjugate for the covariance matrix in the Multinormal process and can easily be seen to be a generalisation of the inverse gamma density function into k dimensions.

$$f_{iW}^{(k)}(\underset{\sim}{S}|\underset{\sim}{\Psi},\nu) = W_k(\nu)^{-1}|\underset{\sim}{\Psi}|^{\frac{1}{2}(\nu+k-1)}|\underset{\sim}{S}^{-1}|^{\frac{1}{2}(\nu-2k)} \exp(-\tfrac{1}{2}\nu \mathrm{tr}(\underset{\sim}{S}^{-1}\underset{\sim}{\psi}))$$

with $$W_k(\nu) = (2/\nu)^{k(\nu+k-1)/2}\, \pi^{k(k-1)/4} \prod_{i=1}^{k} \Gamma\, \tfrac{1}{2}(\nu+k-i)$$

defined for $\underset{\sim}{S}$ positive definite and symmetric and $\nu > 0$.

mean $(\underset{\sim}{S}) = \underset{\sim}{\psi}\, \nu/(\nu-2)$

mode $(\underset{\sim}{S}) = \underset{\sim}{\psi}\, \nu/(\nu+2)$

The parameter ν represents the effective imaginary sample size and HFS and EPS methods are evidently feasible. It is probably easier however to consider the marginal inverse gamma distributions for the diagonal variances and then assess the matrix of intercorrelation coefficients.

WIDER ISSUES

The main concern of this section has been the examination of the structure of various subjective probability assessment procedures, their properties and characteristic biases. The wider social and political aspects of the problem have not been considered. The lack of attention to issues of subject motivation and orientation, social-psychological factors in the decision-analyst and probability assessor relationship, personal resistance to ambiguity and uncertainty, and training programmes is not to be interpreted as a relegation of the undoubted crucial importance of these factors for successful implementation, but a recognition of the fact that many of these issues are common to the practice of operational research and not essentially structural. It should be realised, however, that the practical implementation of such highly formalised subjective techniques as this is open to even more abuse than usual. Phillips

(22) discusses many of the important factors in the assessment task which are necessary in persuading a reluctant subject to respond and in minimising the bias which is introduced by the very presence and behaviour of the analyst himself.

It is argued in Bunn (6) that the general context of policy analysis requires the adoption of a forecasting approach to subjective probability estimation. Thus, there is often an inevitable separation of the probability and utility assessment tasks in the orthodox decision analysis. As a consequence of this, however, because the probability assessors are now in a non-motivating, decision-neutral state, extra care has been taken to minimise causal bias (recall the discussion on the generalisability of laboratory evidence upon anchoring and conservatism bias).

When subjective probability forecasts are produced on a repetitive, team basis, a good example are the U.S. weather forecasters, the use of penalty functions or scoring rules (c.f. Winkler & Murphy (37)) have been successful. Essentially, an error function is defined upon the actual result and the forecaster's probabilistic assessment such that in attempting to maximise his score, he will be increasing the accuracy of his estimation.

Another consequnce of this concentration upon encoding structures is the partial elimination from detailed consideration of the fundamental cognitive processes whereby individuals formulate their beliefs and perceptions. Very little is known about these processes and much of the research must necessarily fall within the scope of neuropsychology. The basic topic, however, is one of aggregating information and constructing sensible inferences. At a normative level, the research on hierarchical inference (c.f. Peters (21)) and on the combination of forecasts (c.f. Bunn (4)) can provide the basis for formal procedures, or at least conceptual structures, in meeting this need. The more explicit an individual's reasoning, the easier it is to identify the precise points of disagreement within a group and hence achieve a consensus. Developments along this line of explicating the 'thinking algorithms' underlying an individual's subjective probability assessment are at present restricted by the lack of a suitable notational logic and representation but are evidently in the spirit of subjective probability in attempting to make statistics less, paradoxically, subjective in the sense that the subjectivity in the analysis is more clearly defined. It should be recognised, however, that control over an individual's formulation of probability judgement can never attain the level of deriving a completely 'unbiased' estimate without the probability ceasing to be subjective.

Consensus and Expert Resolution

These aspects are dealt with at length in the paper by Harman and Press and there is no need therefore for this section to provide the same extensive review as previously given to probability assessment. The brief review given here is therefore rather more by way of a motivating preview.

The approaches to the problem of consensus and expert resolution can be divided essentially into those which are synthetic, i.e. which provide a procedure which a decision-maker may use to synthesise a set of opinions, or those which represent a true consensus insofar as the members of the group reconcile their options equally amongst themselves. In the former case, the methodology is the same as if each member of the group were a forecasting model; the problem is only one of defining and revising the appropriate weights for each individual. True consensus, on the other hand, treats each individual as an equal decision-maker with the implication to represent the best compromise of all options. A synthetic consensus represents the best compromise only from the point of view of the aggregating decision-maker. Thus, the methodologies of true consensus are more akin to democratic voting, Paretian optimality, etc., than the combined forecasts approach of synthetic consensus.

The simplest synthetic consensus will consist of subjectively assessed probabilities on the part of the decision-maker. Winkler (36) and Morris (19), for example, present Bayesian methods for assessing and revising these probabilities. The synthesised consensus, C_s, is then evaluated as the linear "expectation".

$$C_s = \Sigma p_i \, z_i \qquad \forall \; i$$

where z_i is the opinion of the ith member and p_i the associated probability.

Such procedures must generally assume that the experts are in some sense "independent" which given similarities in training and experience, may not generally be valid.

Methods of true consensus would attempt to derive these linear weights by some intrinsic procedure. The paper by Harman and Press is an example of using a simultaneous equation approach to this end.

Structuring and Optimisation under Multiple Conflicting Objectives

Considerable attention is given in this volume to the problem of dealing with multiple conflicting objectives. Baecher, Gros and McCusker discuss various methods for dealing with multiple trade-offs. Brooks applies the multiattribute approach of Raiffa and Keeney in a case study on hazardous shipment decisions. Byer and de Neufville develop a screening procedure aimed particularly at simplifying this problem.

In the paper by Aubin and Naslund, emphasis is placed upon a different aspect of the problem. It is a mathematical programming approach which provides an efficient search procedure by which the decision-maker can identify his most preferred decision. It does not involve the explicit assessment of a multiattribute utility function.

Even after an efficient screening of the policy options, the final set may still contain too many possibilities to be evaluated by a simple decision tree. This is where the need for the mathematical programming methods arises. Furthermore, many policy decisions involve discrete outcomes. The paper by Zionts describes a method which introduced integer programming into the multiple criteria problem, and is therefore addressed specifically to this problem.

Further aspects involved in structuring and assessment are highlighted in the applications papers of Tremolieres and Warner-North, Offensend and Smart. The latter of these provides a detailed case-study illustrating some of the difficulties in applying a cost-effectiveness approach in public policy analysis. Wider issues in the implementation of formal analytic methods for policy analysis are discussed in the paper by Stringer. He particuarly emphasises the organisational aspects of policy formulation which often pose the most extreme constraints and appear, as yet, to have received scant attention from decision analysts in particular and operational research in general.

References

1) Alpert and Raiffa, (1969) A Progress Report on the Training of Probability Assessors. Unpublished Working Paper, Harvard.

2) Barclay, S. and Peterson, C.R. (1974) Two Methods for Assessing Probability Distributions. Technical Report 73-1.

3) Broyles, J.E. and Thomas, H. (1973) Financial Risk in Portfolio Selection. London Business School Working Paper.

4) Bunn, D.W. (1975) A Bayesian Approach to the Linear Combination of Forecasts. Opl. Res. Q.

5) Bunn, D.W. (1975) Anchoring Bias in the Assessment of Subjective Probability. Opl. Res. Q.

6) Bunn, D.W. (1977) Policy Analytic Implications for a Theory of Prediction and Decision. Policy Sciences.

7) Bunn, D.W. and Thomas, H. (1973) Assessing Subjective Probability in Decision Analysis. NATO Conference, Luxembourg and published by White, D.J. and Bowen, K. Decision Theories in Practice, Hodder and Stoughton.

8) du Charme, W.M. and Peterson (1968) Intuitive Inference and Normally Distributed Populations. J. Expt. Psychol.

9) Eilon, S. and Tilley, R.R. (1975) Stochastic Dominance for Ranking Ventures. Omega.

10) Good, I.J. (1965) The Estimation of Probabilities. MIT Press

11) Hadar, J. and Russell, W.R. (1959) Rules for Ordinary Uncertain Prospects. Amer. Econ. Review.

12) Hampton, J., Moore, P.G., and Thomas, H. (1973) Subjective Probability and its Measurement. Journ. Roy. Stat. Soc.

13) Hanoch, J. and Levy, H. (1969) The Efficiency Analysis of Choices Involving Risk. Rev. Econ. Studies.

14) Hill, M. (1973) Planning for Multiple Objectives. Monograph 5 Regional Science Research Institute, Philadelphia.

15) Lavalle, I.H. (1970) An Introduction to Probability, Decision and Inference. Holt, Rinehart and Winston.

16) Markowitz, H.M. (1954) Portfolio Selection. Wiley, New York.

17) Moore, P.G. and Thomas, H. (1975) Measuring Uncertainty. Omega.

18) Morrison, D.G. (1967) Critique of Ranking Procedures in Subjective Probability Assessment. Mgmt. Sci.

19) Morris, P.A. (1974) Decision Analysis Expert Use. Mgmt. Sci.

20) von Neumann, J. and Morgenstern, O. (1947) Theory of Games and Economic Behaviour. Princeton.

21) Peterson, C. (1974) Hierarchical Inference. Decisions and Designs, Inc.

22) Phillips, L.D. (1973) The Psychology of Measuring Probability. London Business School Decision Analysis Unit. Working Paper.

23) Phillips, L.D., and Edwards, W. (1966) Conservatism in a Simple Probability Inference Task. J. Expt. Psychol.

24) Phillips, L.D., Hays & Edwards, W. (1966) Conservatism in Complex Probabilistic Inference. IEEE Human Factors in Electronics.

25) Phillips, L.D., and Thomas, H. (1973) Assessing Probability and Utility. London Business School Decision Analysis Unit. Working Paper.

26) Pickhardt, R.L. and Wallace, J.B. (1974) A Study of the Performance of Subjective Probability Assessors. Decision Sciences.

27) Raiffa, H. and Schlaifer, R. (1961) Applied Statistical Decision Analysis. Harvard.

28) Smith, L.E. (1967) Ranking Procedures and Subjective Probability Distributions. Mgmt. Sci.

29) Stael von Holstein, C.A. (1972) Probabilistic Forecasting. Organisation Behaviour and Human Performance.

30) Thomas, H. (1973) The Assessment of Project Worth with Applications to Research and Development, in 'Cost Benefit and Cost Effectiveness', Allen and Unwin London.

31) Tversky, A. (1974) Assessing Uncertainty. J.R. Statist. Soc. B.

32) Tversky, A., and Kahnemann, D. (1973) Judgement Under Uncertainty. 4th Research Conference on Subjective Probability, Utility and Decision-Making. Rome. (Published in Science, 1974).

33) Whitmore, G.A. (1970) Third Degree Stochastic Dominance. Amer. Econ. Rev.

34) Williams, D.J. (1969) A Study of a Decision Model for R & D Project Selection. Opl. Res. Q.

35) Winkler, R. (1967) The Assessment of Prior Distributions in Bayesian Analysis. JASA.

36) Winkler, R. (1968) The Consensus of Subjective Probability Distribution. Mgmt. Sci.

37) Winkler, R. and Murphy, A. (1973) Subjective Probability Forecasting in the Real World. 4th Research Conference on Subjective Probability, Utility and Decision Making. Rome.

38) Youssef, Z.I. and Peterson, C. (1974) Intuitive Cascaded Inferences Organisational Behaviour and Human Performance.

PHILIP BYER
RICHARD de NEUFVILLE

CHOOSING THE DIMENSIONS AND UNCERTAINTIES OF AN EVALUATION

The procedure developed in Chapter 5 presents pragmatic guidelines, based upon decision analysis concepts, for screening the assumptions involved in a policy analysis. The use of screening coefficients is presented and illustrated in an application to seismic building codes.

5. Choosing the Dimensions and Uncertainties of an Evaluation
by Philip Byer and Richard de Neufville

Introduction

Most real problems requiring policy formulation are complex, featuring many possible alternatives and numerous interest groups with multiple conflicting objectives. In addition, the preferences of these groups are generally nonlinear, and the future performance of each alternative is almost always uncertain. Analysts have tended, for simplicity, not to consider many of these complicating factors in their evaluations of alternative policies. This has often led to analyses, such as the Roskill Commission's benefit-cost evaluation of the Third London Airport, that failed to capture essential aspects of the problem; whose recommendations were frequently rejected; and that were, on balance ineffective. This phenomenon has spurred analysts to incorporate the more realistic but complicating factors into the evaluation process, particularly the multiple interest groups and multiple objectives. For example, environmentalists in the United States have pressured the Federal Government to consider environmental as well as economic and other factors in making decisions about policies that affect the environment.

EVALUATION TECHNIQUES

Promising new techniques are being developed to help analysts include more complex features in the evaluation process (5). Social cost-benefit analysis, for example, attempts to incorporate nonlinear preferences. It values each attribute on the basis of people's willingness to pay for them, as revealed by demand functions, and then combines these values to arrive at a single measure of the net benefits of each alternative (8). This measure of demand purportedly represents the value of each alternative to society. This was the method used to evaluate alternative sites for a Third London Airport (2).

Multiobjective analysis, a procedure now required for essentially all water resources projects in the United States, represents a different approach. This technique defines a surface representing the dominant possible combinations of achievement for each objective, that is, the production possibility frontier. The slopes of this surface show the tradeoffs, or rate of transformation, between the objectives. Given a separately defined possibility nonlinear, utility function, the analyst can identify the alternative that maximises this utility (7).

Multiattribute decision analysis adds other complexities to an evaluation. By incorporating both probabilities and utility, it explicitly consider uncertainties, nonlinear preferences and multiple objectives (9). Multiattribute decision analysis has proven to be both useful and practical (3).

THE NEED FOR ASSUMPTIONS

Whatever form of evaluation is chosen, it is too difficult, too costly or too time-consuming to consider all of a problem's real characteristics. Although some techniques can incorporate multiple attributes into the evaluation process, the number of attributes must be greatly limited due to the difficulty of defining preferences over more than a few dimensions. In addition, since much of the effort in an evaluation is spent on predicting impacts, analysts can spare themselves much work by disregarding an attribute. Furthermore, if the level of an attribute is assumed to be known with certainty, then the analyst avoids the trouble of estimating its probability distribution for each alternative.

Since the benefits of making simplifying assumptions about the problem can be considerable, analysts almost always make them. In doing so, they face a major question: what simplifying assumptions should be made? For example, should certain attributes be included in the analysis or is it reasonable to neglect them? When is it desirable to assume certainty, rather than uncertainty, about the levels of an attribute? The answers to such questions have significant consequences for the cost and practicality of any analysis.

The choice of assumptions should be made carefully. While we may be tempted to neglect certain objectives and uncertainties for the sake of simplicity, we may in so doing discard some of the turning points of the issues at hand. This could lead us into trouble with unacceptable recommendations. The evaluation of alternative sites for the Third London Airport is a case in point. In great part due to its disregard of important land use considerations, as Buchanan pointed out, the recommended solution was ultimately unacceptable. Analysts can also run into trouble by not explicitly considering uncertainties. The evaluation of strategies for the development of nuclear power in Britain in the 1960's for instance, apparently failed to consider the riskiness of innovative programs and, thus, led to a program which placed all effort on a single technology with little provision for exploiting more effective designs (that did ultimately come forth).

The choice of assumptions about a problem also has important implications for the choice of evaluation techniques. Since different techniques incorporate different assumptions into the evaluation, the choice of assumptions largely determines which techniques are most appropriate. For example, cost-benefit analysis does not explicitly account for uncertainties and is only appropriate if certainty and nonlinear preferences are assumed. Similarly, if we assume decision-makers have linear preferences for a single objective, then it may be appropriate to use a simple benefit-cost analysis comparing the expected values of the alternatives. In this case, the use of a linear utility function under uncertainty is equivalent to assuming the expected values of the attribute with certainty. A multiattribute decision analysis however, would be needed if non-linear preferences, uncertainties, and multiple objectives seem important. Byer (1) discusses the choice of evaluation techniques based upon making appropriate assumptions about the problem.

CHOICE OF ASSUMPTIONS

The decision to make an analysis simpler, though less realistic, should be based upon a priori estimates of the significance of added complexity to the effectiveness of the analysis. If it appears that a simpler assumption will not materially affect the choice of optimal policy, then it is reasonable to simplify the analysis by making that assumption.

Different sets of attributes, for example, do not always imply different optimal policies. The rankings of alternative policies based upon two sets of attributes ($\underset{\sim}{x}$, Y) and ($\underset{\sim}{x}$) may be identical either because the level of the attribute Y does not vary sufficiently over the alternatives, or because the decision-maker does not value it highly enough. Similarly, the rankings generated by assuming certainty and uncertainty about the level of an attribute may be identical because of the shape of the utility function and the shapes of the underlying probability distributions over the attribute. As an extreme example, if the probability distributions indicate that the level of some attribute Z is nearly deterministic for each alternative, then it is reasonable to assume with certainty, that the alternatives will result in these "deterministic" values.

Even if a different optimal policy with a lower value would result using a simplifying assumption, this difference may not be significant. If the perceived difference in the values of the optimal policies is less than the limits of the accuracy of the analysis, then we would be unsure which of the two policies would actually turn out to be the best. In this case the simplifying assumption should probably be made. Even if the difference in the values of the optimal policies is significant, the analyst may save more in the costs of the evaluation by making the simplifying assumption than is lost in this difference.

Therefore, while we should recognise the many dimensions of a problem and its essential uncertainties, it may not be worthwhile to insist upon these realities in practice. The choice of assumptions should depend upon the degree and value of the expected changes in both that optimal policy and the cost and time savings realised. Our selection of the evaluation procedure should rest upon the tradeoffs between its cost and effectiveness for the various assumptions.

SCREENING MODELS

Procedures, called "screening models", have been developed to investigate these tradeoffs. They are pragmatic techniques that give first-order estimates of the sensitivity of the optimal policy to different assumptions. We use these models right at the start, before the formal evaluation. Their results help us to decide which factors are important; help us screen out the complexities that will have little bearing on the final evaluation. The subsequent detailed analysis then focuses only on what the screening model indicates is likely to be important. These screening procedures gives us confidence that we are not wasting time and money on irrelevant detail; and also that we do include everything that is important.

The most commont type of screening model is aimed at reducing the number of alternative policies to be evaluated. It entails the use of simple techniques, such as linear programming, to evaluate all of the alternatives and screen out the unpromising ones. The remainder are then evaluated in greater detail. A number of cases demonstrate that this approach can be highly effective. By making it possible to examine many alternatives to some degree, instead of exhausting one's resources on a detailed evavluation of a short list, this procedure has led to improvements of 20 to 30 percent on major projects (4).

Two related models, one that screens the number of objectives or attributes for a problem, and one that screens uncertainties about their levels, have been developed by the authors (1). These models are based upon simple numerical criteria requiring only a little information about preferences and probabilities and only a few simple calculations. Basically, the screening model for attributes develops first-order estimates, in terms of one of the attributes, of the maximum expected difference between the true values of the recommended policies that would result from including and excluding the attributes being screened. Similarly, the screening model for uncertainties about the levels of an attribute calculates an estimate of the maximum expected effect, in terms of that attribute, of assuming some certain levels, such as the expected values, rather than the true range of uncertainties.

To screen out attributes and uncertainties, we compare each of these estimates, called a "screening coefficient", to the magnitude of the attribute that measures it. If it is insignificant, such as by being less than the limits of accuracy in measuring the attribute, we presume that the change in ranking that might occur by making the simplifying assumption is also insignificant. Since upper bound estimates are used, it is more appropriate to use the models to argue that certain attributes should be screened out and that certainty should be assumed, rather than that other attributes should be included and uncertainties assumed.

SCREENING COEFFICIENTS

Estimates of the screening coefficients are easily obtained through the use of standard approximations to both utility functions and the marginal (unconditional) probability distributions for each attribute x_i. Utility functions for a set of attributes can be approximated in terms of the utility functions, $u_i(x_i)$, for each attribute (or, more precisely, a subset of the attributes), and of scaling factors, k_i, between the several attributes

$$u(x_1,x_2,\ldots,x_N)= \frac{1}{k} \{ \prod_{i=1}^{N} [k.k_i u_i(x_i) + 1] - 1\} \qquad (1)$$

Table 1: Measures of Values Used to Calculate the Screening Coefficients

Functions Assumed for Each Attribute, x		Parameters Required	Measures of Value of	
Utility	Probability Distribution		Attribute $\bar{u}$	Uncertainty $\lvert \bar{x} - x_e \rvert$
$a + be^{-cx}$	Normal	a, b, c mean, $\bar{x}$ variance, σ^2	$a + be^{(c^2\sigma^2/2)-c\bar{x}}$	$\lvert c\sigma^2/2 \rvert$
	Exponential (shifted and inverted)	a, b, c upper bound, x^* β [mean, $\bar{x}=x^*-\beta$]	$a + b\frac{e^{-cx^*}}{(1-c\beta)}$	$\lvert \beta + \frac{\ln(1-c\beta)}{c} \rvert$
	Gamma (shifted)	a, b, c lower bound, x_* α, β [mean, $\bar{x}=x_*+\alpha\beta$]	$a + b\frac{e^{-cx_*}}{(1+c\beta)^{\alpha}}$	$\lvert \alpha\beta - \alpha\frac{\ln(1+c\beta)}{c} \rvert$
	Uniform (shifted)	a, b, c upper bound, x^* lower bound, x_* [mean, $\bar{x}=(x_*+x^*)/2$]	$a + b\frac{(e^{-cx_*} - e^{-cx^*})}{c(x^* - x_*)}$	$\lvert \frac{x^*+x_*}{2} - \frac{1}{c}\ln\left[\frac{\bar{u}-a}{b}\right] \rvert$
$a + bx$	Any	a, b mean, $\bar{x}$	$a + b\bar{x}$	0

where

$$k = \{ \prod_{i=1}^{N} [k.k_i + 1] \} - 1$$

as developed by Keeney (6).

To use this formula the x_i's must meet certain reasonable assumptions about their relationship to each other. The utility function for each attribute can, in turn, be approximated by the nonlinear function $a_i + b_i e^{-c_i x_i}$ if the decision-maker is risk averse or risk prone, or by the linear function $a_i + b_i x_i$ if he is risk neutral. As the degree of risk aversion or proneness increases, $|c_i|$ increases; neutrality toward taking risks corresponds to $c_i \approx 0$. The base, a, and the scale, b, of the function are set so that it ranges from 0 to 1 over the range of the attribute, as required by the multiattribute function. The scaling factor, k_i, for each attribute depends on the range of the attribute and its relative importance to the decision-maker.

The approximation of a multiattribute utility function, therefore, only requires us to obtain two parameters for each attribute: c_i, measuring the attitudes toward risk; and k_i, measuring its value relative to the other attributes. A first-order estimate of a nonlinear utility function over two dimensions, for instance, requires estimates of only four parameters. Each of these estimates can be obtained from answers to a few questions asked of the decision-maker (1,6,9). The screening models also require that each attribute be defined such that more of the attribute is preferred.

Probability functions can likewise be approximated by obtaining a priori estimates of only the two or three parameters needed to specify a particular member of any of the few important families of distributions that could reasonably represent a situation of interest. The screening coefficients can, thus, be calculated once a handful of parameters have been estimated. Table 1 illustrates what is involved. For any of the combinations of utility and probability density functions shown, it displays the specific parameters that must be estimated and the formulas for the measures of value used to generate a value of a screening coefficient. The column headed $\bar{u}$ contains the formulas for the expected utility associated with any attribute, x. With an exponential utility function, it is a linear transformation of the moment generating function of the probability distribution. It is simply a linear transformation of the mean of the distribution with a linear utility function. This measure provides the basis for judging whether it is worthwhile to include that attribute in the more detailed analysis.

The column labeled $|\bar{x} - x_e|$ contains the equations for the difference that might arise from assuming, with certainty, that the value of x is its expected (mean) value instead of using its entire probability distribution. The $\bar{x}$ is the mean of the distribution over x; and x_e is the level of the attribute whose utility equals the expected utility. This latter quantity is the certainty equivalent of the alternative; and $\bar{x} - x_e$ is then simply the risk premium implicit in the nonlinear utility function. If preferences are truly linear, then this difference vanishes for all alternatives. Figure 1 illustrates the relationships between these values. If we assume the entire probability distribution over x for each alternative, then our choice of optimal policy will be based upon the true expected utilities. If, however, we assume the values of x to be their expected values, then our decision will be based upon our preferences for $\bar{x}$ as measured by $u(\bar{x})$, for example. These measures can also be easily dervied for families of distributions other than those shown in Table 1.

The procedure using these measures is now briefly described. A technical appendix provides additional details, while Byer (1) gives a complete description. To screen any attribute, x_i, we search for the sets of values of the probability parameters, which vary with the alternatives, that maximise and minimise the corresponding expected utilities, $\bar{u}$. This maximum and minimum are labeled $\bar{u}_i^*$ and $\bar{u}_{i*}$. Their estima-

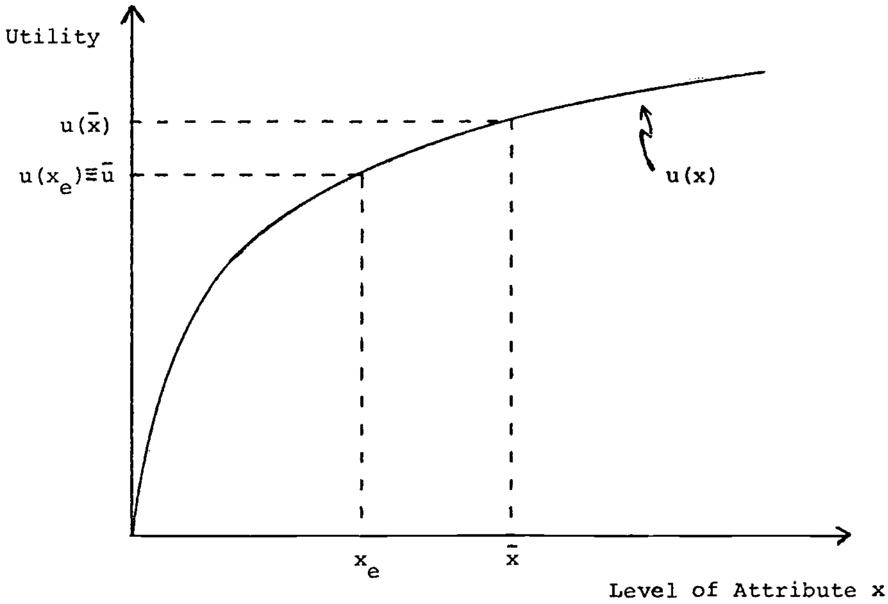


FIGURE 1: TYPICAL SINGLE-ATTRIBUTE UTILITY FUNCTION ILLUSTRATING RELATIONSHIPS USED IN SCREENING UNCERTAINTIES

tion requires us to look at the probabilities of x_i for only a few, rather than all, of the alternatives.

We then choose one of the attributes as a measure of the screening coefficients for all of the attributes. By calculating these coefficients in terms of the same measure, we not only have estimated the absolute importance of each attribute, but also their importance relative to each other. We label this choice x_j. Its screening coefficient is simply the difference between the certainty equivalents corresponding to the maximum and minimum expected utilities $\bar{u}_j^*$ and u_{j*}.

For each of the other attributes, the screening coefficient is given by

$$\max \left\{ \left[\frac{1}{k.k_j}\right] \left[\frac{k.k_i\bar{u}_i^*+1}{k.k_i\bar{u}_{i*}+1} - 1\right],\ \left[1 + \frac{1}{k.k_j}\right] \left[1 - \frac{k.k_i\bar{u}_{i*}+1}{k.k_i\bar{u}_i^*+1}\right] \right\} \Delta x_j \qquad (2)$$

where Δx_j is the range of x_j over which its function is derived.

This equation, which is derived from the multiattribute function, provides an estimate, in terms of x_j, of the maximum expected change in the true value of the recommended alternatives caused by screening out the attribute x_i. Notice that for either a greater scaling factor, k_i, or a greater range of expected utility, $\bar{u}_i$, which reflect the attribute's greater value to the decision-maker or greater variation over the alternatives, the screening coefficient increases. It is then less desirable to screen out this attribute. If however the probability distributions over an attribute do not vary significantly over the alternatives, then the range on the expected utilities would be narrow and, depending upon the value of the attribute to the decision-maker, it may be reasonable to screen out that attribute.

In certain circumstances, we may not easily be able to estimate the bound on the expected utilities for some x_i. For example, if we lack sufficient probabilistic information about an attribute, then it may not be worthwhile to estimate these bounds. In these cases, we can use 0 and 1 as the bounds on the expected utilities, since we have defined each utility function to be between these values. Doing this will give us an upper bound on the value of the screening coefficient.

To screen out uncertainties about the levels of any attribute, x_i, we similarly search for the achievable set of probability parameters that maximises the corresponding risk premium, $|\bar{x}_i - x_{ie}|$. This maximum is the screening coefficient for uncertainties. It is an estimate of the maximum, over all alternatives, of the expected effects of assuming the expected value of the attribute rather than the set of probability distributions.

We can see from the equations for $|\bar{x} - x_e|$ in Table 1 that the magnitude of this screening coefficient depends on the degree of uncertainty, as measured, for example, by the variances of the distributions; and to the degree of risk aversion, as measured by the parameter c. This agrees with out notion that it is more reasonable to assume certainty when the variances are small. Even if they are large, however, it might be reasonable to assume expected values if the utility function is approximately linear ($c \simeq 0$), that is, if the decision-maker is nearly risk neutral. We know that if the decision-maker is truly risk neutral towards an attribute, then expected value and expected utility are the same, in which case it is appropriate to assume its mean value for each alternative.

A further implication of an insignificant maximum $|\bar{x} - x_e|$ arises if x is to be the only attribute in the evaluation. With a single attribute, the alternatives can be ranked according to the relative magnitudes of the certainty equivalents of this attribute. If the screening coefficient indicates that the expected values are never significantly different from the certainty equivalents, then an analysis, such as a simple benefit-cost analysis, based upon linear preferences may be appropriate.

This screening coefficient can also be defined to test the assumption of values other than the means, such as the modes or medians of the distributions. In general, it requires much less effort to estimate the means, modes or medians of the distributions for all alternatives than to estimate the entire distributions.

While screening the coefficients are only first-order estimates, and the decision as to whether they are significant or not is obviously a matter of judgement, it is already clear from the examination of previous studies that the application of these models could have saved substantial effort or led to more effective analyses (3, 10). The following case study illustrates the procedure and its usefulness.

CASE STUDY: SEISMIC BUILDING CODES

The screening models are applied here to the problem of finding the optimal level of resistance to seismic activity that should be required for buildings of a given type and in a given location. This question is now being addressed by the Seismic Design Decision Analysis project in the Department of Civil Engineering at MIT (10). The complete case study is presented by Byer (1). We only present the highlights.

The choice of building code depends on tradeoffs between the additional initial costs of meeting a higher level of seismic resistance and the potential costs, in terms of property damage and human lives and injuries, due to earthquakes.

It has been suggested that the alternative policies should be evaluated on the basis of expected monetary costs, possibly including constant monetary values placed on a fatality or injury. The use of these expected values is tantamount to assuming a linear utility function. The screening models were used to investigate this suggestion. Would that type of evaluation capture the appropriate level of complexity and detail for the analysis of the problem, or is it necessary to consider fatalities, the uncertainties of the consequences, and the nonlinearity of preferences?

The example looks at the evaluation of building codes for 5- to 20-storey, reinforced concrete buildings. Two different kinds of effects are taken to be potentially important: monetary costs and lives lost. The possible design codes span the range from the least protection given by the 1970 US Uniform Building Code to the most. We applied the models separately to two groups, characterised by significantly different preferences, who are concerned with the design: developers and government officials. It was also carried out for two types of locations, seismically high and low risk areas, corresponding approximately to Long Beach, California and Boston, Massachusetts. This definition of the situation allows us to show that different assumptions may be appropriate for different groups and different locations. The numbers used are adapted from data collected as part of the MIT project.

Parameter Estimation

The parameters need to determines the approximate functional representations of the utility functions are listed in Table 2, where the subscripts M and L refer to the dimensions of monetary costs and lives lost. M is the negative of the present increase in present value monetary costs over the initial costs of building without increased protection, assuming a 5% discount rate. It includes initial building costs, measures for increased earthquake resistance, and repair or replacement costs from damage to the buildings due to an earthquake. L is the negative of the percent of building occupants who are killed as a result of earthquakes in the next 50 years.

These parameters indicate that the developers and officials are risk averse toward the monetary attribute, with the latter being much more risk averse. The developers appear to be very risk prone toward fatalities, while the officials would be risk neutral.

To obtain the parameters of the probability distributions of the attributes for a building designed according to any code, we need to multiply the probability of occurr-

TABLE 2: ESTIMATED VALUES OF THE PARAMETERS OF THE UTILITY FUNCTIONS FOR DEVELOPERS AND OFFICIALS

Parameters		Group	
Type	Symbol	Developers	Officials
Risk Aversion Coefficients	c_M	0.001	0.01
	c_L	-0.4	0
Base and Scale Constants	a_M	3.86	1.05
	b_M	-2.86	-0.05
	a_L	0	1.0
	b_L	1.0	+0.05
Scaling Factors	k_M	0.96	0.99
	k_L	0.32	0.99
	k	-0.91	-0.9999

TABLE 3: PROBABILITY OF EARTHQUAKES OF DIFFERENT INTENSITY IN HIGH AND LOW RISK AREAS

Earthquake Intensity (Modified Mercalli Intensity)	Annual Probability Assumed for Risk Area	
	High	Low
≤ V	0	0.975
VI	0.600	0.020
VII	0.350	0.004
VIII	0.045	0.001
IX	0.004	0
X	0.001	0

Table 4: Probability of Effects of Specific Earthquake Intensities on Most and Least Stringent Designs

Design Level	Money Lost (% of Initial Cost)	Lives Lost (% of Total)	Probability of Damage Associated with Modified Mercalli Intensity ≤ V	VI	VII	VIII	IX	X
	0	0	1.00	0.27	0.15			
Least	0.3	0		0.73	0.48			
Stringent,	5.0	0			0.33	0.20		
UBC	30.0	0.25			0.04	0.41		
Zone 0	100.0	1.00				0.34	0.75	0.25
	100.0	20.00				0.05	0.25	0.75
	0	0	1.00	0.67	0.30			
Most	0.3	0		0.33	0.49	0.40	0.10	
Stringent,	5.0	0			0.21	0.52	0.30	
UBC	30.0	0.25				0.08	0.58	
Superzone	100.0	1.00					0.02	0.90
	100.0	20.00						0.10

TABLE 5: ESTIMATED PARAMETER VALUES FOR THE EXPONENTIAL PROBABILITY FUNCTIONS FOR DAMAGE

Design Level	Risk Area	Money x_M^*	Money $-\beta_M$	Lives x_L^**	Lives $-\beta_L$
UBC	High	0	-74.0	0	-5.33
0	Low	0	- 1.25	0	-0.074
UBC	High	-6.7	-16.3	0	-0.22
Superzone	Low	-6.7	- 0.22	0	-0.001

ence of earthquakes of different intensities by the probability of damage to a structure for several levels of possible shaking. These data are given in Tables 3 and 4 and were adapted from those developed by the MIT Seismic Design Decision Analysis effort (10). Only the least and most stringent codes are considered, since they will give us the estimates of the bounds on $\bar{u}$ and $|\bar{x} - x_e|$ required by the screening models.

Shifted, inverted, exponential distributions were fitted to the products of these distributions. The parameters of the estimated distribution of effects appear in Table 5. The x_M* and x_L* represent the upper bounds on these inverted distributions, that is the additional initial costs and no lives lost. Since we define our baseline in measuring costs as the building without any protection, x_M* for the least stringent code is zero by definition. For the most stringent code, x_M* = -6.7 means that the estimated cost of beefing our building up to that level is 6.7% of the basic initial costs. The β_M and β_L are the expected additional effects over 50 years, discounted as indicated above. The mean of the distribution is then the sum of these two quantities.

Significance of the Attributes

We now apply our screening models to help us judge what degree of sophistication is appropriate for which users in what situations. We first turn to the question of what effects it is worthwhile considering in the evaluation of different building code policies. When would we look at both monetary costs and loss of life, and when might be reasonably simplify the analysis by considering only one or the other? To answer this, we calculate the screening coefficient for each of these effects, represented by x_M and x_L.

We first estimate the bounds on the expected utilities. They are calculated by substituting the parameter values in Tables 2 and 5 into the appropriate equation in Table 1. The results bounds appear in Table 6. They show, as should be expected, that the range of expected utilities for each attribute increases as the risk of an earthquake increases. We then choose one of the attributes as the scale on which to measure the screening coefficients. For convenience, we choose x_M.

Finally, we use this information to calculate the screening coefficient, in terms of percent building costs, for each of the two attributes, as explained before. Their values for the different situations appear in Table 7. In viewing them, remember that they reflect the maximum expected effect of omitting each attribute from the formal evaluation. They are, in essence, the difference, in terms of percent of building costs and modified by nonlinear utility functions, between the expected values of the attribute for the two extreme codes. For example, the screening coefficient for money for the developers in the high risk area is 53.8. This is slightly greater than the difference between the expected values over 50 years of 74.0 and 23.0 (= 6.7 + 16.3), as found in Table 5, due to the slight aversion to risks over money by this group. This coefficient for officials is much larger because their greater risk aversion gives enormous weight to high losses, which pushes the perceived amount of loss considerably higher. For low risk areas, where the probabilistic component is much less important, the dominating factor is the 6.7% of initial costs, which is required to prepare the building for the most stringent code.

The magnitudes of the coefficients for fatalities are somewhat more complicated to understand. They, too, are based upon the range of the expected values and attitudes toward risk. However, since they are in terms of the monetary attribute, they also depend on the value of lives relative to money, as expressed by the scaling factors.

The coefficients indicate that monetary losses always constitute an important aspect of the evaluation of seismic codes. For developers, omission of this factor could affect the perceived value of any policy by an amount equal to about 54% of the initial costs of construction in a high risk area, and 6% in a low risk area. These amounts are large both absolutely and relative to potential additional building costs. Equal or greater values apply for officials. Consequently, it appears that monetary costs should, indeed, be part of the formal evaluation.

TABLE 6: ESTIMATED BOUNDS ON THE EXPECTED UTILITIES FOR THE DIFFERENT ATTRIBUTES

Group	Risk Area	Bounds on Expected Utilities			
		$\bar{u}_{M*}$	$\bar{u}_{M}^{*}$	$\bar{u}_{L*}$	$\bar{u}_{L}^{*}$
Developers	High	0.773	0.933	0.319	0.918
	Low	0.980	0.996	0.971	1.000
Officials	High	0.858	0.986	0.733	0.989
	Low	0.996	0.999	0.996	1.000

TABLE 7: SCREENING COEFFICIENTS FOR VARIOUS ATTRIBUTES FOR DIFFERENT GROUPS IN DIFFERENT AREAS

Risk Area	Type of Attribute	Coefficient (% of Initial Basic Costs) for Developers	Officials
High	Money	53.8	110
	Lives	66.0	280
Low	Money	5.67	5.66
	Lives	4.05	85.3

Loss of life, likewise, also appears to be a significant factor in the evaluation. As the likelihood of an earthquake decreases, however, the consideration becomes less important. It may be reasonable to exclude it from the evaluation for developers in a very low risk area. For officials, this exclusion would be warranted only in areas where there is essentially no risk to life, since they value lives quite heavily.

Significance of Uncertainties

We now turn to the question of whether it is desirable, as a practical matter, to assume uncertainties by requiring the use of entire probability distributions. The simpler assumption of using expected values permits us to focus on the means of the distributions rather than to work always with the more complicated distributions.

The screening coefficients for judging the importance of using the distribution rather than just the means are calculated using the probability parameters for the least stringent code, where the most damage is expected, in the appropriate formula in Table 1. We calculate this coefficient for each attribute separately in terms of itself. Table 8 shows the results of these calculations. In viewing them, one should remember that they do not estimate the actual expected effect but, rather, the maximum expected perception of the decision-maker of what it means to neglect all possible values of x other than its mean. The 60.6 for officials for monetary losses in high risk areas implies that the officials can expect to perceive a neglect of these uncertainties to be potentially equivalent, at the maximum, to a loss of about 60% of the initial basic costs of the building. This high value reflects the fact that officials weight calamitous loss very heavily due to the nonlinearity of their preferences.

The screening coefficients indicate that considerable accuracy could be lost in the evaluation if one disregards the inherent uncertainties in the monetary losses in a high risk area. The maximal difference of 60% is large both absolutely and relative to the potential size of the monetary loss in this area.

It appears reasonable, on the other hand, to assume expected values for loss of both money and lives in the low risk area. Working with the mean values in these cases would change our perception of the value of any code so minimally that it would only have a trivial - if any - effect on our ranking of alternatives. In fact, because the officials appear to have linear preferences for fatalities, we see no difference caused by making this assumption for them in any area.

This procedure for determining how detailed and complex the evaluation should be rests on judgment, of course. It is, consequently, not altogether unambiguous. Consider, for example, the screening coefficient for monetary losses for developers in high risk areas. Although a 2.9% increase in costs is reasonably large absolutely, it is small compared to the potential magnitude of the losses, which could be total. The coefficients are an aid to judgement - where no other quantitative measures exist - but not a substitute for judgement.

The screening models lead to several fairly strong conclusions for this case. First, it seems quite clear that neither monetary losses nor fatalities should generally be excluded from the formal evaluation of building codes for either developers or officials. Such an analysis based solely on monetary costs is, therefore, inappropriate.

Another conclusion is that in low risk areas it is reasonable to compare preferences for expected effects, that is expected monetary costs and expected fatalities, to evaluate the alternative codes. A cost-benefit analysis, for example, would seem adequate in this case. A multiattribute decision analysis, however, might be more appropriate in high risk areas.

CONCLUSION

The procedure outlined and illustrated here is a pragmatic guide based on decision analysis concepts, to what kind of evaluation is needed in any situation. It provides a

TABLE 8: SCREENING COEFFICIENTS FOR UNCERTAINTIES

FOR DIFFERENT GROUPS IN DIFFERENT AREAS

Risk Area	Attribute		Coefficient fer	
	Type	As % of	Developers	Officials
High	Money	Initial Basic Building Costs	2.9	60.6
	Lives	Occupants	2.5	0
Low	Money	Initial Basic Building Costs	0.001	0.01
	Lives	Occupants	0.001	0

mechanism for objectively addressing the question of whether particular aspects of a problem, which we know to exist, are worth taking into account in a practical situation. The screening coefficients contain valuable information, albeit first-order estimates, about the effects of various choices of assumptions. The procedure may result in entire categories of effects being screened out, and the indication that other attributes, which may not otherwise be included, are important. As shown in the case study, the models can have important implications for the analysis of policies that have potentially catastrophic consequences.

Further work is being done by the authors to develop screening models for other types of assumptions, such as the linearity, rather than non-linearity, of preferences.

References

1) Byer, P. "A Screening Model for the Selection of Evaluation Techniques for Engineering System Designs," unpublished Ph.D. Thesis, Department of Civil Engineering, Massachusetts Institute of Technology, June 1975.

2) Commission on the Third London Airport, Report, London: Her Majesty's Stationery Office, 1971.

3) de Neufville, R. and R.L. Keeney, "Use of Decision Analysis in Airport Development for Mexico City," Systems Planning and Design: Case Studies in Modeling, Optimisation and Evaluation (see ref. 4), 1974, pp 349-369.

4) de Neufville, R. and D.H. Marks, eds., Systems Planning and Design: Case Studies in Modeling, Optimisation and Evaluation, Englewood Cliffs, N.J.: Prentice-Hall, 1974.

5) de Neufville, R. and J.H. Stafford, Systems Analysis for Engineers and Managers, New York: McGraw-Hill Book Company, 1971.

6) Keeney, R.L., "Multiplicative Utility Functions," Operations Research, Vol. 22, No. 1, Jan.-Feb. 1974, pp 22-34.

7) Marglin, S.A. Public Investment Criteria, Benefit-Cost Analysis for Planned Economic Growth, Cambridge, MA: MIT Press, 1967.

8) Mishan, E.J., Economics for Social Decisions, Elements of Cost-Benefit Analysis, New York: Praeger Publishers, Inc., 1973.

9) Raiffa, H. Decision Analysis, Introductory Lectures on Choices Under Uncertainty, Reading, MA: Addison-Wesley, 1968.

10) Whitman, R.V., J.M. Biggs, J.E. Brennan, III, C.A. Cornell, R. de Neufville and E.H. Vanmarcke, "Seismic Design Decision Analysis," Journal of the Structures Division, American Society of Civil Engineers, Vol. 101, No. ST5, May, 1975, pp. 1067-1084.

TECHNICAL APPENDIX

This section briefly describes the rationale and derivation of the screening models. Complete details are given by Byer (1). First, we identify a set, x, of attributes that measure all potentially important effects, and partition this set into mutually exclusive, collectively exhaustive subsets, x_i, such that these subsets are preferentially and utility independent, as defined by Keeney (6). We can generally reasonably assume, particularly in a screening procedure, that attributes that measure different types of effects meet these conditions.

Keeney has shown that if these conditions are met, then the utility function (of the von Neumann-Morgenstern type defined over risk) over all of the attributes is either an additive or multiplicative function of the utility functions over each of the subsets. The multiplicative form is shown in equation 1. The additive form is

$$u(x_1,x_2\ldots,x_N) \equiv u(x) = \sum_{i=1}^{N} k_i u_i(x_i) \qquad (3)$$

The $u_i(x_i)$'s are utility function over each x_i, and k and the k_i's are constant scaling factors, the levels of which depend upon the range and relative preferences for the attributes.

The utility functions are defined over the range of interest, x_{i**} to x_i^{**}, of the attributes, such that

$$u_i\ (x_{i**}) = 0$$

and (4)

$$u_i\ (x_i^{**}) = 1$$

for all i.

The k_i's are scaling factors for each $u_i(x_i)$ relative to u(x). It is equal to the probability p_i such that the decision-maker is indifferent between receiving $x_{1**},x_{2**},\ldots,x_{i-1**},\ x_i^{**},\ x_{i+1**},\ldots,x_{N**})$ with certainty and receiving $x_1^{**},x_2^{**},\ldots,x_N^{**})$ with probability p_i or $(x_{1**},x_{2**},\ldots x_{N**})$ with probability $1-p_i$. These factors can be estimated through a series of a few questions of the decision-maker.

If the sum of these factors equals unity, then the additive form of u(x) (equation 3) is appropriate. Otherwise, the multiplicative form (equation 1) should be used.

The expected utility of any alternative, D_m, is given by

$$\bar{u}_m(x) \equiv \int_{x_{**}}^{x^{**}} u(x) f_m(x)dx \qquad (5)$$

where $f_m(x)$ is the probability distribution corresponding to this alternative for the occurrence of the possible levels of the set of attributes.

Similarly, the expected utility with respect to some x_i is given by

$$\bar{u}_{im}(x_i) \equiv \int_{x_{i**}}^{x_i^{**}} u_i(x_i) f_{im}(x_i)dx_i \qquad (6)$$

where $f_{im}(x_i)$ is the marginal (unconditional) probability density function over x_i. The level of x_i whose utility is equal to this expected utility is the certainty equivalent, x_{ime}:

$$u_i(x_{ime}) = \bar{u}_{im}\ (x_i) \qquad (7)$$

It can be shown that, with the multiplicative or additive multiattribute utility function, the expected utility of any alternative is bounded by the utility of its certainty equivalent for x_i and its bounds on the other attributes.

$$u(x_{ime}, x_{\bar{i}**}) < \bar{u}_m(x) < u(x_{ime}, x_{\bar{i}}^{**}) \tag{8}$$

where $x_{\bar{i}}$ is the set of all of the attributes except x_i.

We now define the lower and upper bounds, x_{ie*} and x_{ie}^*, on the certainty equivalents over all alternatives

$$u_i(x_{ie*}) \equiv \bar{u}_{i*} = \min_m \left[u_i(x_{ime})\right] \tag{9}$$

and $$u_i(x_{ie}^*) \equiv \bar{u}_i^* = \max_m \left[u_i(x_{ime})\right]$$

then

$$u(x_{ie*}, x_{\bar{i}**}) \leq u_m(x) \leq u(x_{ie}^*, x_{\bar{i}}^{**}) \tag{10}$$

for all alternatives. $\bar{u}_{i*}$ and $\bar{u}_i^*$ are the bounds on the expected utilities over x_i.

For every alternative, there will be some level of $x_{\bar{i}}$, labeled $x_{\bar{i}m}$, such that

$$\bar{u}_m(x) = u(x_{ime}, x_{\bar{i}m}) \tag{11}$$

where

$$x_{ie*} \leq x_{ime} \leq x_{ie}^* \tag{12}$$

and

$$x_{\bar{i}**} \leq x_{im} \leq x_{\bar{i}}^{**} \tag{13}$$

Suppose that we screen x_i out of the evaluation. This would be comparable to ignoring its level or, it can be shown, to assuming that x_{ime} is any value between its bounds shown in equation 12, for every alternative. The maximum expected change caused by screening out x_i, in the value of any alternative, including the optimal one, is, therefore, measured by

$$x_{ie}^* - x_{ie*}.$$

To put the value to the decision-maker of this difference in terms of some other attribute, say x_j, $i \neq j$, we must find the change in x_j that has equal utility to this difference. Figure 2 illustrates the tradeoffs between x_i and x_j that define this equivalence. The curves in this figure are iso-utility (indifference) curves over x_i and x_j. Because of the assumption of preferential independence, their shapes are independent of the levels of the other attributes. For alternative D_m, the utility to the decision-maker of the change from x_{ie}^* to x_{ie*} is equivalent to the utility of changing from x_{jm}'' to x_{jm}'. However, we do not know $u_m(x)$ for any alternative. An upper bound estimate on the difference $x_{jm}'' - x_{jm}'$, which does not require knowing any expected utilities, is the maximum of $x_j^{**} - x_j'$ and $x_j'' - x_{j**}$, where

$$u(x_j'', x_{ie*}, x_{\bar{ij}}) = u(x_{j**}, x_{ie}^*, x_{\bar{ij}}) \tag{14}$$

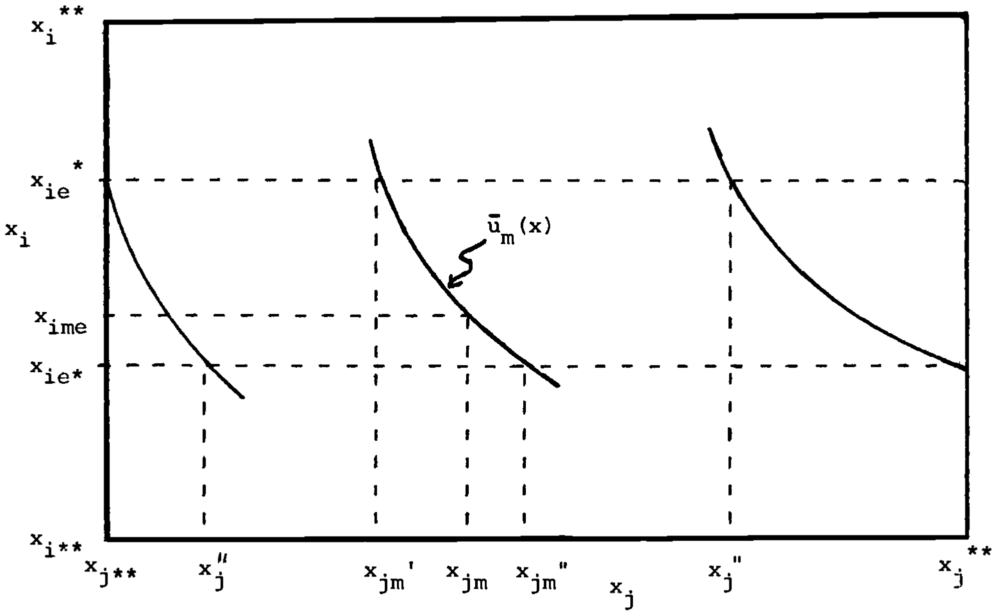


FIGURE 2: ISO-UTILITY CURVES ILLUSTRATING TRADEOFFS USED TO DEFINE THE SCREENING COEFFICIENTS

and

$$u(x_j', x_{ie}^*, x_{\bar{i}\bar{j}}) \quad u(x_j^{**}, x_{ie}, x_{\bar{i}\bar{j}}) \tag{15}$$

where $x_{\bar{i}\bar{j}}$ is some value for the other attributes. These utilities provide reasonable and convenient bounds on the expected utilities.

Using these equalities with the equations for u(x), we find that, if u(x) is of the multiplicative form

$$u_j(x_j'') = \left[\frac{1}{k.k_j}\right] \left[\frac{k.k_i u_i(x_{ie}^*) + 1}{k.k_i u_i(x_{ie*}) + 1}\right] - \frac{1}{k.k_j} \tag{16}$$

$$u_j(x_j') = \left[1 + \frac{1}{k.k_j}\right] \left[\frac{k.k_i u_i(x_{ie*}) + 1}{k.k_i u_i(x_{ie}^*) + 1}\right] - \frac{1}{k.k_j} \tag{17}$$

and, if u(x) is of the additive form

$$u_j(x_j'') = \frac{k_i}{k_j} \left[u_i(x_{ie}^*) - u_i(x_{ie*})\right] \tag{18}$$

$$u_j(x_j') = 1 - \frac{k_i}{k_j} \left[u_i(x_{ie}^*) - u_i(x_{ie*})\right] \tag{19}$$

To find a convenient expression for $x_j'' - x_{j**}$ and $x_j^{**} - x_j'$, for $i \neq j$, we assume $u_j(x_j)$ to be linear

$$u_j(x_j) = \frac{x_j - x_{j**}}{x_j^{**} - x_{j**}} \tag{20}$$

to get

$$x_j'' - x_{j**} = u_j(x_j'') \left[x_j^{**} - x_{j**}\right] \tag{21}$$

and

$$x_j^{**} - x_j' = \left[1 - u_j(x_j')\right] \left[x_j^{**} - x_{j**}\right] \tag{22}$$

where $u_j(x_j'')$ and $u_j(x_j')$ are defined by equations 16 through 19.

The maximum of $x_j'' - x_{j**}$ and $x_j^{**} - x_j'$, defined by equations 21 and 22 and 16 through 19, is the screening coefficient for x_i, $i \neq j$. It is an upper bound measure of the maximum expected change, caused by screening out this attribute, in the true value of the alternative that is recommended. Only the screening coefficient assuming the multiplicative utility function is shown in the text (as equation 2) because of its more general applicability.

These coefficients require estimates of expected utilities over x_i. We can approximate any single attribute utility function, $u_i(x_i)$, with either a linear or exponential function

$$u_i(x_i) = a_i + b_i x_i \qquad (23)$$

or

$$u_i(x_i) = a_i + b_i e^{-c_i x_i} \qquad (24)$$

The expected utility, $\bar{u}_{im}(x_i)$, and their certainty equivalents, x_{ime}, that result from combining these utility functions with univariate probability density functions can often be easily calculated, using general equations that are in terms of the parameters of the utility and probability functions that are assumed.

The expected utility that results from using a linear utility function and any probability density function is given by the utility of the mean of the distribution, which is to say that the mean and certainty equivalent are equal. This corresponds to the bottom line of Table 1.

Using the exponential utility function

$$\bar{u}_{im}(x_i) = \int_{x_{i**}}^{x_i^{**}} \left[a_i + b_i e^{-c_i x_i} \right] f_{im}(x_i) dx_i \qquad (25)$$

$$= a_i + b_i \int_{x_{i**}}^{x_i^{**}} e^{-c_i x_i} f_{im}(x_i) dx_i$$

This last integral is simply the moment generating function, $M[c_i, f_{im}(x_i)]$, of the probability density function. The certainty equivalent is such that

$$a_i + b_i e^{-c_i x_{ime}} = a_i + b_i M[c_i, f_{im}(x_i)]$$

$$x_{ime} = \frac{\ell n M[c_i, f_{im}(x_i)]}{-c_i} \qquad (26)$$

which is the x_e used to obtain the coefficient in the column farthest to the right in Table 1.

ALVIN J. HARMAN
S. JAMES PRESS

ASSESSING TECHNOLOGICAL ADVANCEMENT USING GROUPS OF EXPERTS

This paper presents methods for collecting and analysing judgements from groups of experts. It addresses issues associated with resolving the procedural and administrative problems involved in selecting a panel of experts, in eliciting informed judgements about the degree of technological advance on relevant projects, and in designing a survey questionnaire for measuring those judgements.

Three methods of multivariate analysis are described for quantifying and analysing group judgement data collected from a panel of experts. Those of the methods that are known in earlier literature include multi-dimensional scaling of individual differences, and subjective probability procedures, including the Bayesian approach. A new procedure, which was developed specifically for this application, involves the use of simultaneous equation system models in which the response (dependent) variables are categorical and unordered.

6. Assessing Technological Advancement Using Groups of Experts

*by Alvin J. Harman and S. James Press

This paper is concerned with some methodological aspects of collecting and analysing expert group judgement data. We consider the context of data obtained from experts by survey questionnaire. We examine questions associated with how to design the survey and the survey instrument so that the responses might be most informed, most useful for policy decisions, and most "correct." We then discuss some methods for analysing such data.

This work arose out of considering the problem of how to assess the degree of technological advance some new military "system" might have (and its estimated cost would be expected to be a function of the "ambitiousness" of the system development phase and the system's degree of "sophistication). It was felt that technical "experts" would probably be in the best position to assess the risks of attempting various technological advancements in a given time period for some new system proposed but not yet designed. Thus, illustrations of general principles are typically in the technological advance context.

---

* We are grateful to the National Research Council of Canada and The Rand Corporation for their financial support.

1 *Collection of Expert Group Judgement Data: Survey Design Considerations*

This section focuses on the problems of design of procedures for eliciting and grouping expert judgements and forecasts. It is believed that by eliciting the judgements of experts, and by studying the spectrum of their views on the relevant issues in a formal way (without necessarily looking for or encouraging consensus), one will be able to improve on existing bases for budget allocation which currently rely on a somewhat less formal methodology.[1]

The basis for this approach to assessing technological advance rests with the notion that there is such a thing as expertise about the subject, and that many experts are better than one. This section discusses some of the considerations surrounding these ideas and attempts to establish reasonable criteria for ultimately eliciting a collection of expert judgements on the same set of questions.[2] The basic considerations can be divided into four broad categories -- the existence of expertise; the identification and selection of a panel of experts; the formalism and procedural issues associated with eliciting responses; and the design of a meaningful measuring instrument (a survey questionnaire). These problems are discussed, in turn, below.

1.1 *Existence of Expertise*

In many problems it is hard to argue that there is such a thing as expertise. For example, suppose it is six months before a national election and the question is, "who will win the race?" No one really knows and the degree of knowledge in the hands of people who make a career out of studying elections is not significantly different, at this time, from that of the average person. In another context, suppose we are interested in speculating about the "qualities of everyday life" in the year 2000 A.D. It is difficult to imagine that there is a greater degree of knowledge, intuitive understanding and ability to predict such "qualities" in the hands of some few people compared with the rest of us. (In fact, if such greater knowledge did exist it is quite unclear as to how it would be identified -- but that is yet another kind of problem, and one which is considered below).

The basic idea behind expertise is that for some problems there exist people who really have so much more knowledge and understanding of the mechanisms underlying the phenomena in question that they can do an appreciably better job of forecasting long term trends and changes than a relative "layman" (thus, a layman is a non-expert). The notion of visiting the oracle at Delphi to receive "expert" advice is an old one. People sought Delphic advice on the complete range of human questions and problems, in spite of whether or not expertise really existed. We make the same mistake today of thinking that for every problem there exists an expert problem solver.

In some problems, there is no doubt that expertise does exist. A good physician can do a better job of assessing the likelihood of cancer developing in a given individual than a layman; a good lawyer can generally do a better job of assessing the likely behaviour of a judge or jury, in a given context, than a layman; and a good scientist or

---

[1] One mechanism for eliciting the judgements of experts on "fuzzy" issues and then grouping them ot obtain a consensus has been the Delphi Technique. The survey methods to be discussed below differ in many important respects from the conventional Delphi approach.

[2] We are not addressing the often important problem of pooling the judgement of a group of people about "values" (such as judgements about relative "importance" of items in a set, relative "desirability" of various items, and relative "goodness" of various behaviour patterns). In such problems there really are not "experts" in the knowledge sense, but there may be a certain subpopulation whose judgements are more relevant than others (such as the member of the Board of Directors of a corporation that must make a decision about values).

engineer can do a better job of forecasting technological change than a person who is not technically trained or experienced.

In the context of an R&D program, once the planning objectives are carefully defined, the first step in an evaluation is to stand back and to take a hard look at the question of whether or not expertise really exists on the feasibility of various technological advances that would contribute to these objectives. If the answer is negative, we must seek an alternate path for evaluation. If the answer is positive, we can proceed to the next set of considerations.

1.2 *Identification and Selection of a Panel of Experts*

In the sequel we assume that expertise exists for the problem at hand but now it is a question of picking a panel of experts. Many considerations are involved: How do we recognise an expert? How heavily should each view be weighted? How many experts constitute a "good" panel? Should all experts have the same type of expertise? How many panelists with each type of exertise should there be? Do the experts believe they are experts and how does that affect their judgement? What are the common characteristics of experts and, in selecting a panel, should their qualities be matched? These are some of the complex issues which must be addressed.

Attributes of a Good Panel of Experts

Several characteristics seemed to be important for panels charged with assessing potential developments associated with technological change:

(1) *Diversity*

The panel members reflect a wide spectrum of talents. That is, good panels should probably not be monolithic in terms of the field of expertise represented. Rather, for every aspect of the problem under study, there should probably be some panel member who is expert in that area. This characteristic diversity of disciplines represented is necessary in order that the panel avoid overlooking or giving perfunctory treatment to fundamentally important facets of the problem.

(2) *Depth*

Some panel members should have a profound understanding of the technical issues involved. They should be considerably more knowledgeable, in a scientific sense, than most people in the world, in their particular speciality. For every major scientific area which is a component of the basic problem there should be at least one expert with great depth in his subject.

(3) *Breadth*

Good panels should probably contain some members who are "systems experts." That is, there should be some individuals who are accustomed to thinking on a broad level -- e.g., in terms of the interactions of various subsystems; in terms of the implications of new subsystem developments on economic feasibility of an entire system; and in terms of political, legal, social, and ecological overtones of the new development. Panel members who have this type of breadth of knowledge are probably better able to predict feasibility and likelihood of large technological development taking place than the layman, who in this case may be some "deeply knowledgeable" scientific expert who tends to be quite narrow in his views and who tends to ignore other developments which will be needed to render developments in his own field meaningful.

It is not clear what mix of experts is most appropriate on a panel. What fraction of the members should be system people and what fraction discipline experts? We can at least establish lower gounds, however. That is, once the problem has been broken down into some well defined fields in which expertise exists, we believe there probably ought to be at least one expert from each field, and at least one systems analyst whose breadth has burgeoned out of that field.

While our assertions about what constitutes a "good" panel of experts have not been substantiated here, they are empirically verifiable, and we expect that experiments will validate our conclusions.

Identification of Experts

A reasonable definition of scientific expertise involves recognition and approbation of peer groups; to wit, someone is an expert in his field if others in his field consider him to be an expert. Some measures of expertise, by this definition, are the holding of office in the national scientific organisation, the holding of a position on the editorial board of the important technical journals in the field, awards for outstanding scholarship, honorary positions in national societies, publications of non-introductory books (publications of monographs and advanced treatises), and the holding of awards of research contracts from various branches of the federal government. In many situations involving scientific expertise, such measures taken jointly, rather than singly, would very likely serve as useful identifiers of expertise. When a variety of professionals in a field are polled about whom they regard as an expert, and the same individuals keep being mentioned, those individuals must be considered experts.

Outstanding systems analysts have typically been technical experts at one time and then later chose some type of administrative path of personal development. Their perspective has broadened and their knowledge of related fields has perhaps decreaeed. They began interacting more with known experts in each of the fields required for the analysis, and increasingly found instances where the dominating constraint on a development involved some field other than their own. After coping with many diverse developmental efforts these individuals increasingly found themselves able to predict feasibility, timing, and likely constraints associated with any new technological construct. These individuals are currently employed as some type of manager (academic department chairman, research director, R&D manager for a corporation or a governmental agency, etc.).

The immense value of having people with the above characteristics present on the panel stems not only from their broad prospective, but also from the fact that they tend to counterbalance the very conservative viewpoints typically found amoung scientific experts. That is, individuals with a deep knowledge of a scientific subject have spent many years being indoctrinated to exercise extreme caution (if not suspicion) about scientific breakthroughs and meaningful technological advances. Such people are not inclined to admit that they jump to conclusions, and so as a group they tend to be conservative about the feasibility, timing, and costs of new developments. Such a posture is "safe", given their elevated status. If they are wrong, it won't be in the absurd direction, so it is unlikely they will then be subject to criticism, ridicule, and loss of status. The systems analysts tend to be less conservative, ridi- often halve the time estimates for a new development given to them by a scientific and expert on their staff.

Selection of Panel Members

To minimise selection bias, panelists should be chosen by standard procedures developed in Statistics and the Theory of Psychological Measurement.[1] It is easy to see how careless selection methods could reflect institutional rivalries and personal biases of the people who do the selection. What is needed to start with is an exhaustive listing of *all* known experts in each of the fields required for the analysis,

1. American Psychological Association (1966).

and a similar listing for the "systems" people.[1] This will establish populations of experts. Then, after stratifying by field, a collection of simple random samples might be taken from each list (population). One alternative might be to stratify still further by preparing lists, for each field, which give experts in government, in industry, and in the academic world. Then we might choose a simple random sample from each of these categories, in each field. Systems analysts should also been chosen for the panel by stratified ramdom sampling. The procedures in the two cases are completely analogous. It is anticipated that by using random number tables, in the usual way, to choose a ransom sample from these population lists, judgements which are representative of those of the entire lists will be obtained, and results would be fairly repeatable if the survey were carried out on several more occasions with similarly chosen samples. Moreover, the viewpoints or biases characteristic of a certain class of expertise (e.g., originating in industry) can be separately investigated and appropriate allowances made. For this purpose, specific background information should be elicited from each respondent; this information can be explicitly treated within models to analyse the panelists' assessments as described in Section 2 below.

Respondent Motivation

A very important issue associated with panel selection concerns the motivation of the panel members to participate fully in the study. Suppose we have two experts in the same field who, for our purposes, are equivalent in expertise and we wish to compare their responses to a given question. If one of the two experts gives an "off-the-top-of-the-head" response (a response based upon a few seconds or minutes thinking and intuiting), while the other expert takes the time to think through all the steps necessary to reach the final goal or development and evaluate the problems and constraints associated with each state, assessing conditional probabilities for each of the stages, it seems reasonable that we should weight the judgement which was more carefully arrived at, more heavily.[2] Thus, if the panelists are not strongly motivated to cooperate fully, to the extent of providing careful, introspective responses, the results of the survey will not be reliable. Of course we can query the panelists formally on the questionnaire, as to how much time they devoted to preparing their responses; and we can word some of the technical questions so that the panelists are required to provide step-by-step responses as often as possible; but these approaches, while helpful, provide only partial relief from the problem and don't really come to grips with the sources of the difficulty, namely lack of motivation for the panelist.

Various means might be used to induce experts to respond cooperatively (assuming they agree to participate in the study in the first place). Possible motivational techniques include:

(1) *Honoraria*

The payment of a nominal fee to participants. Such a fee could hardly be less than $50 or $100. If there were 30 panel participants, this would imply an honorarium cost of $1,500 to $3,000. But such a token fee is too small to really represent any kind of real inducement to cogitate; more likely, it might be an inducement to some to agree to cogitate; more likely, it might be an inducement to some to agree to participate in the study and to provide merely $50 worth of responses. A significant

---

1. It is certainly conceivable that it may be desirable to stratify the populations by levels of expertise. On the other hand, it is a rare question for which one-person's assessment is singularly appropriate.

2. It is always appropriate to determine whether the respondent has reflected upon this question previously. If so, then a rapid response might be every bit as valid as one arrived at after protracted introspection or analysis.

fee of say $1,000 or more might bring this cost of the study to $30,000 or more, which could be prohibitive. The larger fee, however, is much more likely to induce motivation than a simple honorarium (clearly a $50,000 fee would induce most scientists of the world to be very conscientious; but how much less would do the same job?)

(2) *Participant Listing*

The participants might be promised that they would be listed as a group, after the study is completed, in the written accounts of the study results. Such a approach is in keeping with the notion that seeing one's name in print is a strong inducement to an individual to make sure that the work associated with his name is substantial and correct. But the resulting report will not have the status of a paper in a prestigious scientific journal, nor will the individual be spotlighted as any kind of individual innovator, but will only be listed as one of a group whose opinion was sought. The resulting effect of participant listing on motivation is not likely to be significant, and might be perverse if some experts were concerned that the group's views would not adequately reflect their own.

(3) *Real-time interaction*

The procedures by which information is elicited may influence the quality of that information. Thus, a telephone interview by a technically competent inquisitor can both assess the seriousness with which the panelist is considering the questions and probe to ensure that certain subtleties of the questions have been comprehended. However the perspective of the inquisitor himself may unduly influence the panelist. An alternative is real-time responses and interactions among the panelists via a distributed computer system on which individuals would probe for the reasons behind each others assessments as well as for characteristics of the group of responses. This procedure might tend to encourage group self-motivation without the personal interactions often attributed to in-person panel discussions.

(4) *Research contract award potential*

If the request to participate in the study comes from a potential source of research contract funds, the inducements to many individuals to cooperate in the study may be very great. This would be particularly true if all questionnaires were name-tagged and it was clear to every participant that the survey monitors were keeping track of how each individual was responding. Such a motivation may also lead to biases in responses, as noted above.

1.3 *Procedural Issues*

Once an appropriate panel of respondents has been selected we must be concerned with some procedural and administrative issues. Should the questions of interest be asked in personal interviews with continual interaction and feedback between interviewer and interviewee? Perhaps the questions should be administered by telephone, or by mail; there are advantages and disadvantages to each of these approaches. Perhaps some questions should be addressed with panel members unknown to one another (so that their judgements cannot be impugned on an authoritarian basis), and perhaps there is another group of questions which should be addressed in group discussions with all panel members freely interacting with one another, airing their views openly. Should there be one fixed set of questions or should the questions proceed in stages on one basis or another? Should panelists be asked the same questions repeatedly, after telling them the opinions of other respondents?

How can we ensure that the questions are valid; that is, are they phrased in such a way that they are really providing the answers to the questions we want answered, with minimum semantical difficulties, and maximum focus on the true points of interest? The time honoured method which appears best for checking validity is to use the response results for forecasting and to compare them with actual outcomes. But when we are in a forecasting context in the first place, this approach may be difficult to implement. A pilot study involving short term prediction might prove helpful, but it is unlikely to if the type of expertise involved in short and longer term assessments differ.

Which of the above types of administrative approach should be used to elicit opinions from a panel of scientific experts? Surely the correct answer here depends very much upon the type of question being asked and upon the use to which the responses will ultimately be put. It seems reasonable to expect that questions involving detailed scientific knowledge and expertise are best answered individually, with a minimum of outside bias from other panelists, from an interviewer, or from supervisory personnel. Other questions, such as those which involve the potential use of policy variables which might greatly affect rates of technological change and development, are probably best handled by the group as a whole (although not necessarily in face-to-face discussion). For example, if a researcher knew that his particular work, while apparently not very important in and of itself, was in fact the major limiting factor to an extremely important development, and if he also knew that because of its importance, the funding level of support for his research might be increased 100 per cent, his judgement about the feasibility and timing of some theoretical future development might be drastically changed. This type of background information, while difficult to supply in a questionnaire, (since we can't always anticipate all the implicit questions asked, and the underlying assumptions made, by a respondent), is quickly requested and supplied in a group discussion.

A mail survey instrument might be administered by preceding its mailing with letters and or telephone calls advising that the questionnaire is going to be mailed, and then following up the mailing with telephone calls[1] in the nature of clarification of questions and checking on receipt of the questionnaires. Such a procedure should not only provide more valid questions and should minimise non-response, but also it should maximise respondent introspection and cooperation.

Questions might be asked once, or they might be repeated in stages in a controlled way. There are many advantages to controlled feedback. For example by requiring every respondent to provide some discussion (for example, a paragraph of prose) about why he believes in his first round position, and then later, why he is either adhering to his first round position, or why he is changing from his first round position (after having been supplied with summary information relating to the first round responses from all respondents). In this way, we are forcing all respondents to think through their judgements very carefully, vis-a-vis all other respondents, and we are focusing in on the best rationale for the group judgements. Certainly no type of consensus judgement should be required for this type of analysis,[2] and, in fact, lack of consensus may reflect the degree of compexity of the issue involved.

Questions involving a paragraph of prose of an enumeration of reasons are useful for collecting ideas about how to regard an issue. We believe questions of this type are probably the most reasonable for the first stage of such a study. Later stages might involve more highly focused questions to the same panel.

---

1. In the basic mailing, subjects could be advised that there would be a telephone followup.

2. Controlled feedback is used in the Delphi approach also. However, there it is typically used by panelist to justify their positions only if they are outside the interquartile range of the distribution of responses. Such an approach of course encourages agreement even when it may not be appropriate.

1.4 *Instrument Design*

The design of a suitable instrument for measuring the judgements of scientific experts about the feasibility of some technological developments depends of course to a great extent upon the specific developments of interest, the breadth and depth of the type of information sought, and the degree of detail required. However, prior to actual design, one can at least set down some guidelines and general considerations which should prove useful.

Questionnaire design is an art about which much has been written.[1] There have been many guiding principles set down on the basis of past experience. They include the importance of giving the instrument a preliminary trial run in a pilot program, the care that must be exercised in question wording, and the close attention that must be paid to the design of scales for recording judgements with quantitative content. We will not comment further on these important, but fairly standard, problems associated with all questionnaire designs.

Types of Questions

It seems appropriate, in the class of problem exemplified by the technological advance question, for there to be at least three distinct classes of questions in the instrument:

1) Questions dealing with the backgrounds of the individual respondents;

2) Questions dealing with the nature, format, administration, and execution of the questionnaire;

3) Questions dealing with the scientific content which motivates the entire study

The first category of questions relates to the degree of expertise of the respondent, his biases, the nature of his background (for example, is he a systems manager) and generally his qualifications for being on the panel. There might also be some questions which are aimed at assessing the degree of scientific conservatism of the respondent. Perhaps asking for his judgement on a key scientific question, for which a spectrum of viewpoints is already known, would prove to be a useful procedure for assessing conservatism relative to other panelists. The biases of a respondent might be brought out by asking for a listing of his best publications. The answers to this groups of questions should prove useful for understanding and interpreting the quality and perspective of each individual's responses.

The second group of questions attempts to assess the care that was exercised in completing the questionnaire, whether or not the questions were clear, how the question format might be improved in the next round, and whether or not the respondent feels he was too constrained in his responses by the way in which the questionnaire was administered.

Goals of Questions

The goal of the questions might be to assist a manager in allocating his R&D budget (or more generally to aid in the allocation of scarce resources). In such cases, it is desirable to condition the questions, whenever possible, on policy issues. For example, in asking respondents to assess feasibility of some new development, they might be asked for three separate assessments; one assuming a "low" funding level for

---

1. See, for example, Oppenheim (1966).

the necessary research, a second assuming a "medium" funding level, and a third assuming a "high" funding level (of course, these terms need to be suitably defined either as part of the questionnaire, or as a part of the requisite response). It might be worthwhile to ask for likelihoods of a given development within say 2 years, 5 years, 10 years, or perhaps "never". It would also be useful to ask panelists what other areas of research need to be "stimulated" because of their likely complementary payoffs. These other areas might only relate to peripheral aspects of their own work, or they might broaden the applicability of their work.

Question Format

An important problem in this type of study is how to phrase the questions so that the respondent is led by the questions themselves to reflect thoughtfully about the problems and their constraints and limitations.

One set of questions might permit the respondent infinite latitude by asking him to enumerate all the steps which will be involved in order to attain a given technological development. Another set of questions, which might constrain the respondent somewhat more, would provide the respondent with some basic steps required to attain a given development, but would ask him to add or delete steps, as appropriate, and to assess conditional probabilities of being able to proceed down the chain of steps at each stage. The end result would be an "achievement tree" with many nodal points and assessments of the conditional probabilities of moving between any two nodal points. This degree of detail could be refined in successive stages of the questioning, as could the subjective probability assessments. Moreover, successive stages of questioning could easily lead to the breaking out of completely new paths of development.

2. *Quantification and Analysis of Expert Group Judgement Data*

Research on quantitative assessment of technological advancement by the use of expert judgements appears to have had its formal genesis in studies carried out by Marshall and Meckling (1959), Klein (1962), and Summers (1965), who each made use of a quantity "A", the degree of technological advancement sought in a program. To estimate A, a sample survey was taken with four "experienced Rand Corporation engineers" being used as sample elements. The four subjects were asked to rate subjectively the magnitude of the improvement in the state of the art required for each of 22 aircraft and missile development programs. Their ratings were to be placed on a numerical scale ranging from 1 to 4. Group judgements were assessed by summing the ratings of the four experts. After all the ratings were obtained, each program was categorised as "small", "medium", or "large", in the Marshall and Meckling and Klein studies. The A variable (henceforth called the A-factor), and others, were related to cost factors (ratios of actual to estimated costs) of a program by use of standard regression techniques in the Summers study.

Subjective assessments of the A-factor were attempted in two subsequent surveys. One was reported on by Harman and Henrichsen (1970). Respondents were asked to assess the A-factor for aircraft and missile systems on a scale of 0 to 20. As in the earlier survey, the subjects were experienced Rand engineers, and the sample size for any given system ranged from two to four subjects.

The last of the three surveys was undertaken in 1970 in connection with the aircraft turbine engine.[1] In this survey A-factors were assessed (among other things) for aircraft turbine engines on a scale of 1 to 20. The eleven subjects were all employees of the General Electric Company (some of the systems evaluated were manufactured by G.E. and some were not). Non-response was large and not all subjects made assessments for all systems.

---

1. Alexander and Nelson (1972).

Although the three surveys undertaken to assess A-factors have contributed considerably toward our understanding of the problem of how to measure technological advancement, they still leave much to be desired. For example, samples have been so small that it is difficult to make meaningful statistical statements about the results. Also there are problems associated with asking individuals (no matter how expert) to compare objects having many characteristics or attributes on a single numerical scale. Not only individual perceptions of reality but also individual weights assigned to each attribute of an object being studied will tend to be difficult. Finally, individuals differ in their ability to quantify their judgements; even though they might view some object in the same way, they might very well differ in their quantitative description of it.

This section discusses several approaches to solving the problem of assessing technological advancement by quantifying sets of judgements. These approaches include: a type of multidimensional scaling called "individual differences scaling," subjective probability assessment techniques, and multivariate regression with categorical dependent variables.

2.1 *Multidimensional Scaling of Individual Differences*

A very powerful and by now well known method of integrating collections of comparative judgements of individuals to form a composite group judgement, scaled on each of several dimensions, is called "individual differences scaling."[1] The basic idea, applied in our context, is that each subject compares N projects regarding their relative degrees of technological advancement.[2] Thus, each individual renders $N(N - 1)/2$ judgements of the form: project S_i is more technologically advanced than project S_j, for all $i, j = 1, 2, \ldots, N$. In another context, each individual might order proposed R&D projects according to their probability of feasible development to a given stage by a given date. Next it is assumed that p dimensions are sufficient to represent the structure underlying the project differences. The ordinal judgements for a given individual might now be represented as ranks, or they might be converted into "distances" by one of several standard procedures, such as by the "law of comparative judgement". The distances may now be represented as weighted distances in Euclidian space. The weights on each axis and the coordinates of each point may now be estimated by the data. Thus, if the distance between project i and project j, as perceived by subject k, is

$$d_{ij}^{(k)} = \left[\sum_{t=1}^{p} w_{kt} (X_{it} - X_{jt})^2 \right]^{1/2},$$

$i, j = 1, \ldots, N$, and $k = 1, \ldots, n$, if there are n subjects who render complete sets of judgements. The w_{kt}'s and the X_{it}'s are estimated from the data. The result is a single composite (for the n subjects) configuration of points in p-space representing the relative positions of the N systems. The coordinates are the scale values (on a ratio scale) of each of the projects in each dimension. Thus, a collection of pairwise ordinal rankings for each of n individuals would not only yield a composite set of ratio-scaled numerical values in p-dimensions for each project, but also a set of weights, for each person in the sample, representing the importance that person

---

1. Carroll and Chang (1970).

2. The definition of technological advancement to be used simultaneously for past, current, and future projects has not yet been made precise.

3. See, for example, Torgerson (1958); Bock and Jones (1968, Chapter 8).

places on each axis for each project.

The response of each subject[1] may be thought of as being composed of an overall common mean response, plus a response effect reflecting his particular degree of expertise, plus a white-noise error term accounting for individual variations in ability to express one's state of knowledge (thus, two individuals with precisely the same state of knowledge would still differ in their responses because of their error terms). One implication of this type of reasoning is that since experts should produce responses that are closer to being correct than non-experts, great care should be exercised in selecting the subjects. Once a panel of potential subjects is screened for expertise, choosing a sample from this panel, which should be as large as possible consistent with cost and praticality considerations, will result in a small "individual-difference error".

The results of an individual-difference scaling type of analysis might be used in several ways. One way would be to establish a functional relationship (by regression methods) between cost and the various dimensions of the projects. Then, cost of a new project could be predicted by interpolation (extrapolation). Another use of the results might include changing the objectives sought in the new project if it is found that coordinate values along certain axes are unreasonable or intolerable. Finally, after studying the sets of estimated weighting factors for each expert, it might be decided that certain individuals should be queried regarding their weighting of a particular axis. The result could be the uncovering of important circumstances related to the R&D projects that were overlooked (or ignored) by other subjects . Perhaps subjects should be informed of the weights placed on each axis by others and then the entire process repeated. A fruitful area of research should be simulation of such a feedback process of scaling judgements.

Suppose w_j denotes the px1 vector of weights determined by the individual differences scaling algorithm. Then we can define a covariance matrix

$$S = (n-1)^{-1} \sum_{1}^{n} (w_j - \bar{w})(w_j - \bar{w})'.$$

The quantity, trace (S), is a measure of degree of consensus of the subjects' judgements. Thus, perfect consensus is reached when trace (S) = 0.

2.2 *Subjective Probability Methods*

The degree to which some goal may be achieved can be "scored" on some appropriate scale, or alternatively, the probability of achieving the goal may be assessed. Which approach is best? We now consider this question in the context of technological advance.

A-Factors Versus Probabilities

Technological advancement might be measured in terms of A-factors (the degree of technological advance sought in future programs), as described above, or it might be measured in terms of probabilities of some proposition. For example, if we speak of E_i as the proposition that project S_i will have, upon completion, A-factor A_i, $i = 1, \ldots, N$, then $p_i \equiv P\{E_i\}$ should be a monotonic function of the A-factor. That is knowing p_i is equivalent to knowing the corresponding A-factor, and conversely. The real implication of this equivalence is that if the cost of a project, for example, is to reflect the degree of its technological sophistication, either the A-factor or p_i could be used as an explanatory variable in the cost equations. It is not clear at this time which of the two is a better measure, in the sense that it can be better assessed, and can, therefore, be used to generate better cost predictions. This

---

1. Regression models relating response of each subject to explanatory variables are currently being built and will be reported on at a later date.

point will be considered further below. The relationship between the p_i's and the A-factors is monotonic, but not unique; there are a large number of potentially useful and convenient monotonic functions. For example, a linear relationship is provided by

$$p_i = \frac{A_i - a}{b - a}, \quad b > a,$$

where p_i denotes the probability that S_i will have A-factor A_i, and A_i is scaled on the interval $|a,b|$, where a denotes the minimum degree of advancement and b denotes the maximum. This relationship is depicted in Figure 1.

Another potentially useful functional relationship is the logistic correspondence,[1] given by

$$p_i \doteq \alpha \left[\frac{1}{\beta + e^{-(A_i - a)}} - \frac{1}{\beta + 1} \right],$$

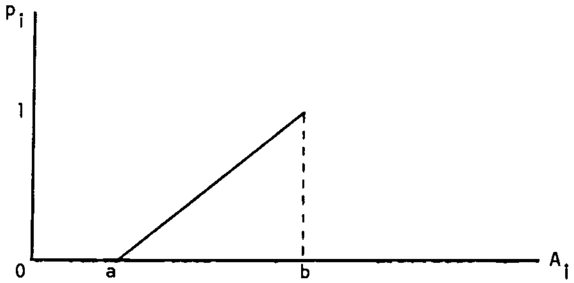


Fig. 1. Linear correspondence between A-factors and probabilities

where $0 < \beta < 1$,

$$a^{-1} = \frac{1}{\beta + e^{-(b-a)}} - \frac{1}{\beta + 1},$$

$a \leq A \leq b$, and $a < b$. In this relationship, p_i is still a monotonically increasing function of A_i, but the function is convex or concave depending upon whether $a < A_i < a - \log \beta$, or $a - \log \beta < A_i \leq b$, respectively. That is, there is a point of inflection at $A_i^* \equiv a - \log \beta$. The functional relationship is sketched in Fig. 2. The appropriate value of β might be selected with the use of regression techniques, after both A-factor and probability information is elicited.

---

1. A multidimensional version of this form of correspondence is used explicitly, in a different context, in Section 2.3.

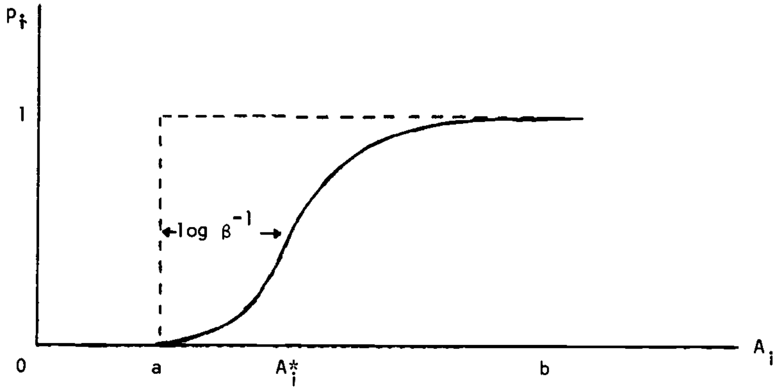


Fig.2. Logistic correspondence between A-factors and probabilities

Many factors enter into the problem of deciding whether probabilities or A-factors are better measures of technological advancement (in the estimation sense described above), some of which are these:

1. One may be more *easily assessable* than the other in that it may be easier to get subjects to render judgements of one rather than the other.

2. One may be more *precisely assessable* than the other in terms of repeatability of such assessment using different samples of "equally competent" experts; that is, it may be subject to smaller variance.

3. One may be a more *accurate* measure than the other in terms of how close to the "truth" the measure can be expected to come (on the average) for a given sample size of independent and identically distributed observations.

4. One measure may be more *improvable* than the other in that it may be possible, by feeding back information to experts over time, to teach them to be better assessors of one measure than the other.

5. Individuals may vary less in their *ability to quantify* their judgements about linearly scaled variables. There is also the difficulty that some people have more of a grasp of the meaning and concept underlying a probability than others. These people have a keener ability to make probability assessments of their judgements than others who might possess the same substantive information regarding a given proposition but are unable to quantify it as well.

Considerable research effort has been devoted to finding methods for getting subjects to assess scores for a variable, a difficult problem that has consumed much time among psychologists and statisticians.[1] The problem of assessing subjective probability

1. See, for example, W.S. Torgerson (1958); Bock and Jones (1968); Guildford (1959); and Coombs (1964).

has received considerable attention not only in psychological and economic contexts,[1] but also in medical contexts.[2] However, whether probabilities or A-factors should be assessed remains an open question.

Bayesian Approach

The Bayesian approach to statistical inference and decision-making involves the assessment of prior distributions on the underlying parameters of the model.[3] One of the most important practical problems associated with the specific application of Bayesian analysis is that of how best to assess the prior distributions. For example, subjective probability assessors are very often "incoherent" in that their probability assessments for various related events are not consistent. For example, for some integer random variable E, an assessor might assert that as far as he is concerned, *a priori*, $P\{E > 0\} = 0.60$ and $P\{E > 10\} = 0.70$. One implication of these two assertions is that $P\{1 \leq E \leq 10\} = -0.10$, an absurd result.

It is expected that computers will be able to assist in the assessment problem. That is, routines could be developed to keep track of all previous assessments an individual has made about related propositions. The computer could ask the subject a sequence of questions designed to lead to consistent assessments of various propositions. If the responses are inconsistent, the computer would indicate this and request the subject to be more introspective and rethink his collective responses. Complete prior distributions could be assessed in this way. This type of computer-assisted assessment technique is a very realistic, potentially available development.[4]

Controlled Feedback Methods

Technological advance might be assessed by some controlled feedback method (the Delphi method is one such approach), which seeks to obtain a composite judgement of a group of experts by feeding individual opinions back to the group to permit the members to revise their assessments. Each expert is typically subjected to a series of questionnaires. The summary statistics of the responses and perhaps the justification for individual's responses are fed back to the panelists, leading to a new round of revised responses. The feedback process is repeated several times. It is anticipated that if consensus is achieved, it will be achieved after several rounds. The experts are generally unknown to one another, and their opinions are often solicited by mail.

The Delphi method was originally devised in 1951 to apply expert opinion to the selection, from the viewpoint of a Soviet strategic planner, of an optimal U.S. industrial target system and to the estimation of the number of A-bombs required to reduce the

---

1. A brief bibliography would include Edwards (1954a, 1954b); Kyburg and Smokler (1964); Preston and Baratta (1948); Mosteller and Nogee (1951); Winkler (1967, (1971); Winkler and Murphy (1968); Einhorn (1972); Savage (1971). For further references see especially Winkler (1971) and Savage (1971).

2. See, for example, Meehl (1974) and Coppelson et al. (1970).

3. See, for example, D.V. Lindley (1965) for discussion of Bayesian inference in univaritate analysis, and S.J. Press (1972) for its application in multivariate analysis. Some earlier work on the assessment of subjective probabilities, in the context of weapons development and the relation of subsystem properties to total system performance, may be found in F.S. Timson (1968).

4. Some computer-assisted assessment techniques are already in use by M. Novick at the American College Testing Program, Iowa City, Iowa.

munitions output by a prescribed amount.[1] Since then the original Delphi techniques and many variations of them have been applied (and mis-applied) to a diverse collection of problems ranging from technological forecasting for advance planning for corporations to studying national goals for the United States and for various foreign countries. Some of these applications are *not* ones for which expertise obviously exists (see Section 2.1).

After study and application of the techniques of the years, a collection of four summarising reports appeared at Rand,[2] and others have reported on recent experimental results.[3] A critique of Delphi which focused on the misuse of controlled feedback has been completed recently.[4]

The idea behind Delphi and other controlled feedback techniques is that if you want the best guess about a "fuzzy" question -- one that is extremely difficult to answer even for the most informed people -- ask an expert. Moreover, since for certain problems many expert heads are better than one, ask many experts and combine the conclusions, weighting them by the degree of expertise in the subject. The details of how to implement this type of philiosophy have varied from one application to another, but the basic idea remains the same. In some applications, the median response is fed back to the subjects for comparison with their own responses, and then a second-round response is sought. In other applications, upper and lower quartiles of the responses, as well as the median, are fed back. In some applications, subjects whose responses fall outside the upper and lower quartiles are requested to explain by they are outliers; otherwise, they are required to change their positions. In still other applications, subjects are required to extrapolate the future from earlier data. Finally, subjects might be required to provide paragraphs of prose describing their feelings, beliefs, or reasoning on an issues. It is the *last* mentioned approach which we believe to be most appropriate, for reasons summarised in Section 2.

2.3 *Categorical-Dependent-Variable Multivariate Regression*

Motivation

This section presents a simplified version of a new methodological procedure for doing multivariate regression analysis, using categorical dependent variables.[5] Moreover, it is shown how the technique can be applied to the problem of assessing technological advance and of comparing the feasibility of competing R&D projects. The methodology generalises the results of categorical-dependent-variable regression, single-equation systems, to correlated, multiple-equation systems of the same form. An extensive treatment of the subject is presented elsewhere.[6]

Suppose there is a panel of n subjects, each member of which is asked to judge technological advancement for N projects, $S_1, \ldots, S_N$. Define the endogenous (dependent) indicator values $y_{ij}(k)$, as follows:

---

1. See Dalkey and Helmer (1951, 1962).

2. Dalkey (1969); Brown, Cochran, and Dalkey (1969); Dalkey, Brown, and Cochran (1969, 1970).

3. Dalkey and Rourke (1971), and Dalkey and Brown (1971).

4. Sackman (1974).

5. Categorical variables can assume only a finite number of values. For example, "dummy variables" that are zero or one, depending upon whether or not some event occurs, are binary categorical variables.

6. Nerlove and Press (1973).

$$y_{ij}(k) = \begin{cases} 1, & \text{if event } E_{ij}(k) \text{ occurs,} \\ 0, & \text{others,} \end{cases}$$

where $E_{ij}(k)$ denotes the event that in a pairwise comparison, the k^{th} subject assesses the degree of technological advancement of S_i to be greater than that of S_j: $i, j = 1, \ldots, N$; $i \neq j$; $k = 1, \ldots, n$.

Next suppose the exogenous (independent) variables $X_{1ij}, X_{2ij}, \ldots, X_{rij}$ bear directly upon whether $S_i > S_j$ (project S_i is more technologically advanced that S_j). Let $x_{1ij}(k), x_{2ij}(k), \ldots, x_{rij}(k)$ denote the values perceived for $X_{1ij}, \ldots, X_{rij}$ by the k^{th} subject. These values really form the basis for the k^{th} subject's assessments and may be thought of as quantitiatively defining the set of assumptions he makes when he renders his comparative judgements.[1]

Adopt the model

$$y_{ij}(k) = F[\beta_{0ij} + \beta_{1ij}\, x_{1ij}(k) + \ldots + \beta_{rij}\, x_{rij}(k)] + u_{ij}(k),$$

where $u_{ij}(k)$ denotes a random disturbance term with the properties that $E[u_{ij}(k)] = 0$ and $E[u_{ij}(k_1)\, u_{ij}(k_2)] = 0$, $k_1 \neq k_2$; and $F(.)$ denotes a monotonic non-decreasing transformation with $0 \leq F(.) \leq 1$; $j = 1, \ldots, N$; $k = 1, \ldots, n$.

For algebraic simplicity, define the r-dimensional column vectors

$$\beta_{ij} \equiv (\beta_{0ij}, \ldots, \beta_{rij})',$$

and

$$x_{ij}(k) \equiv [1, x_{1ij}(k), \ldots, x_{rij}(k)]',$$

so that the model becomes

$$y_{ij}(k) = F[x'_{ij}(k)\, \beta_{ij}] + u_{ij}(k);$$

for $i, j = 1, 2, \ldots, N$; $k = 1, \ldots, n$. In any particular problem we assume that $x_{ij}(k)$ is given (subject k may specify it in part or in toto), and then $y_{ij}(k)$ is generated by assessment (when subject k renders his judgement about S_i vis-a-vis S_j).

Define

$$P_{ij}(k) \equiv P[y_{ij}(k) = 1] = P[E_{ij}(k)]$$

That is, since

$$P_{ij}(k) = E[y_{ij}(k)],$$

$$P_{ij}(k) = F[x'_{ij}(k)\beta_{ij}],$$

$i, j = 1, \ldots, N$; $k = 1, \ldots, n$. Thus, by using the sample of n subjects to estimate β_{ij} as $\hat{\beta}_{ij}$, p^*_{ij}, the probability that $S_i > S_j$, given any preassigned set of assumptions, x^*_{ij}, is estimated as

---

1. For example, the x's may measure the degree of information the k^{th} subject has about the relative difficulty between two projects, or the depth of background or experience the k^{th} subject has relative to projects i and j.

$$\hat{p}^*_{ij} \equiv F[x^*_{ij};\ \hat{\beta}_{ij}] \quad ; \qquad i, j = 1, \ldots, N.$$

But the disturbance terms $u_{ij}(k)$ are mutually correlated for various i and j, for a fixed k (even though they are assumed uncorrelated for different k's), so the $y_{ij}(k)$ are mutually correlated for fixed k. Hence, there is information in one equation that can be used, in part, to estimate parameters in other equations. That is, the system of equations should be viewed simultaneously as a set of multivariate non-linear regression equations in which the endogenous variables (the y's) are discrete, and in which there is systematic heteroscedasticity.[1] The solution to this problem is a set of estimates of p^*_{ij}, for i, j = 1, ..., N, telling the analyst the relative degrees of technological advance required for a set of R&D projects (useful, for example, in cost equations) or telling the policymaker how to view the level of difficulty of a new project. The statistical method devised for solving this problem is outlined below. The less technically inclined reader could skip the remainder of this section without losing any information related to the substantive (non-statistical) issues involved.

Statistical Model

To provide a thumbnail sketch of this approach, the categorical dependent-variable regression model to be used for evaluating technological advancement is illustrated below for the simplified case of a two-equation system in which both endogenous variables are dichtomous (binary). The background, the generalisation to more than two equations are to dependent variables with more than two possible values (polytomous variables), and the discussion of problems of inference, in the general case, are elsewhere.[2]

Let Y_1 and Y_2 denote two binary variables taking on the values zero and one, and suppose Y_1 and Y_2 are correlated.[3] Adopt the two-equation model

$$Y_1 = \theta_1 + \varepsilon_1,$$

$$Y_2 = \theta_2 + \varepsilon_2,$$

where $\theta_1 = E(Y_1)$ and $\theta_2 = E(Y_2)$ are non-random, ε_1 and ε_2 are random disturbances, $E(\varepsilon_1) = E(\varepsilon_2) = 0$, and $E(\varepsilon_1\ \varepsilon_2) = \theta_3$. Note that since $0 \leq \theta_1 \leq 1$ and $0 \leq \theta_2 \leq 1$ and since the variables of Y_1 and Y_2 change with their means (as in any Bernoulli distribution), this is not the usual kind of regression and requires quite a different approach. Since Y_1 and Y_2 can each be zero and one, there are four possible combinations of values. Represent them in a bivariate contingency table:

Y_1 \ Y_2	1	0
1	p_{11}	p_{10}
0	p_{01}	p_{00}

1. Unequal variances of the disturbance terms for a given k.
2. Nerlove and Press (1973).
3. Suppose $Y_j = 1$ if development j will occur by 1983, and $Y_j = 0$ otherwise; j = 1, 2. Both developments might depend upon the same external factors and might dictate that Y_1 and Y_2 be highly correlated.

The entries in the table are the joint probabilities

$$p_{ij} \equiv P[Y_1 = i, Y_2 = j]; \qquad i, j = 0, 1.$$

Note that

$$\theta_1 = E(Y_1) = P[Y_1 = 1] = p_{11} + p_{10},$$

$$\theta_2 = E(Y_2) = P[Y_2 = 1] = p_{11} + p_{01},$$

$$E(Y_1Y_2) = \theta_1\theta_2 + E(\varepsilon_1\varepsilon_2) \equiv \theta_1\theta_2 + \theta_3,$$

$$\theta_3 = E(\varepsilon_1\varepsilon_2) = p_{11}p_{00} - p_{10}p_{01}$$

$$p_{11} + p_{10} + p_{01} + p_{00} = 1.$$

Thus, the problem can be formulated and parameterized in terms of either $(\theta_1, \theta_2, \theta_3)$, or (p_{00}, p_{01}, p_{10}), with $p_{11} \equiv 1 - p_{00} - p_{01} - p_{10}$.

Suppose the four probabilities (only three are linearly independent) characterising the distribution underlying the two-equation system are each related to r exogenous variables, $x_1, \ldots, x_r$, in the same way but with differeing sets of weights.[1] Specifically, adopt the model

$$p_1^* = p_{00}, \; p_2^* = p_{01}, \; p_3^* = p_{10}, \; p_4^* = p_{11},$$

where the p_j^*'s are related to the x_j's through the logistic transformation

$$p_j^* = \frac{\exp(x'\gamma_j)}{\sum_{i=1}^{4} \exp(x'\boldsymbol{\gamma}_i)}; \quad j = 1, \ldots, 4,$$

where γ_j: r x 1 denotes a vector of unknown coefficients and $\Sigma_1^4 p_j^* = 1$. Note that using this type of transformation insures that $0 \leq \theta_j \leq 1$, for $j = 1, 2, 3$. It will be seen that the two-equation dichotomous-variable system that has been placed into the framework of a bivariate contingency table has in turn been replaced by a bivariate polytomous-variable problem in which the dependent variable (it is now a bivariate vector) has four possible values. The coefficient vectors $\gamma_1, \ldots, \gamma_4$ are estimated by maximum likelihood in the following way.

Suppose Y_1 and Y_2 are each observed n times independently, yielding observations $y_i(1)$, and $y_i(2)$, for $i = 1, \ldots, n$. Define the bivariate column vectors

$$y_i = [y_i(1), y_i(2)]', \quad i = 1, \ldots, n.$$

and the constant vectors

$$a_1 = \binom{1}{1}, \; a_2 = \binom{0}{1}, \; a_3 = \binom{1}{0}, \; a_4 = \binom{0}{0}.$$

---

1. If $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ are in turn functions of given exogenous variables $x_1, \ldots, x_r$, we can try to estimate the coefficients of the two relationships and use them for predictive purposes.

Next define the indicator variables

$$v_{ij} = \begin{cases} 1, & \text{if } y_i = a_j \\ 0, & \text{otherwise} \end{cases}$$

Let x_{ij} denote the j^{th} exogenous variable corresponding to the i^{th} observation vector y_i, $i = 1, \ldots, n$, $j = 1, \ldots, r$. Let z_i denote the r-dimensional column vector $z_i = [x_{i1}, \ldots, x_{ir}]'$, $i = 1, \ldots, n$.

Define

$$p^*_{ij} = \frac{\exp(z_i'\gamma_j)}{\sum_{k=1}^{4} \exp(z_i'\gamma_k)},$$

for $i = 1, \ldots, n$, $j = 1 \ldots, 4$. The likelihood function (given $z_i, \ldots, z_n$) is

$$L = \prod_{i=1}^{n} (p^*_{i1})^{v_{i1}} \ldots (p^*_{i4})^{v_{i4}},$$

where $\sum_{j=1}^{4} v_{ij} = 1$. Note that since the p^*_{ij}'s are linearly dependent with respect to j only (r-1) of the γ_j's are required for the problem. Impose the constraint that $\sum_{j=1}^{4} \gamma_j = 0$. Maximising L (actually log L) with respect to $\gamma_1, \ldots, \gamma_4$, subject to this constraint (cf. Press, 1972, p 270), gives the results that the maxmium likelihood estimators $\hat{\gamma}_j$ must satisfy the system of equations.

$$\sum_{i=1}^{n} \left(\frac{e^{z_i' \hat{\gamma}_\ell}}{\sum_{k=1}^{4} e^{z_i' \hat{\gamma}_k}} \right) z_i = \sum_{i=1}^{n} z_i \, v_{i\ell}, \quad \text{for} \quad \ell = 1, \ldots, 4$$

with $\sum_{j=1}^{4} \hat{\gamma}_j = 0$. For predictive purposes, the four cell probabilities in the contingency table are estimated by

$$\hat{p}^*_j = \frac{\exp(x'\hat{\gamma}_j)}{\sum_{i=1}^{4} \exp(x'\hat{\gamma}_i)}, \qquad j = 1, \ldots, 4.$$

A computer program for estimating the parameter vectors is currently operational and in use at Rand, University of British Columbia, Northwestern Unviersity and the University of Chicago.

It is often of interest to study whether or not Y_1 and Y_2 are independent. Since for a given x, Y_1 and Y_2 are independent if and only if $p_{11}p_{00} = p_{10}p_{01}$, the implication is that they are independent if and only if

$$\gamma_1 - \gamma_2 - \gamma_3 + \gamma_4 = 0.$$

Hence, to investigate independence, it is only necessary to estimate the γ_i's and then examine the term $(\hat{\gamma}_1 - \hat{\gamma}_2 - \hat{\gamma}_4 + \hat{\gamma}_4)$ to see if it is significantly different from zero. Since in large samples the γ_u's are normally distributed (they are maximum likelihood estimators), the linear combination of interest is also normally distributed and tests of significance are straightforward.

Multiequation systems (more than two) can be studied in the same manner as above using multidimensional contingency tables, increasing the order of the tables for non-binary variables. The result will be a multivariate polytomous dependent variable to be analysed as above.

Interpretation

Suppose, for example, there are three systems to be compared regarding technological advancement. In the section on Motivation, $y_{ij}(k)$ was defined to be one or zero, depending upon whether or not in a pairwise comparison the k^{th} subject assesses the degree of technological advancement of S_i to be greater than that of S_j, $i, j = 1, \ldots, N$, $i \neq j$, $k = 1, \ldots, n$. Take $N = 3$ (three projects to be compared) and define the random variable Y_{ij} for which $y_{ij}(k)$ is the k^{th} observed value. The three random variables, Y_{12}, Y_{13}, Y_{23}, are mutually correlated and completely describe the comparative states of advancement of the three systems (note that if ties are ruled out, $Y_{ij} = 1 - Y_{ij}$ so that random variables Y_{ij} for which $j < i$ are unnecessary). Now rename the variables, $Y_1 \equiv Y_{12}$, $Y_2 \equiv Y_{13}$, $Y_3 \equiv Y_{23}$. Then, taking $Y_i = \Theta_i + \varepsilon j$, $j = 1, 2, 3$, where ε_i denotes an error term with mean zero, puts the problem into a three equation model (rather than the two equation model described in this section). The j's will of course be taken to be the monotone transformations of linear combinations of independent variables discussed above. In this form the system can be thought of as a 2^3 or 2x2x2 (trivariate) contingency table which can, in turn, be thought of as a trivariate categorical dependent variable regression equation in which the dependent variable can assume eight possible values. If N systems are to be compared there will be $N(N - 1)/2$ simultaneous equations to be solved in this way, rather than the three used in the example.

3. *Discussion and Summary*

We have seen how diverse procedures developed in different disciplines might be brought to bear on the problem of how to combine the opinions of individuals to form a group judgement about an ill-defined, multidimensional concept, such as the degree of technological advance required to complete a given R&D project, or the probability that a certain technological development will become feasible by a preassigned date. Although there are advantages with each procedure suggested, there are also various difficulties, uncertainties, and limitations, both conceptually and technically.

In summary, the approaches of individual-differences scaling, subjective probability methods, and categorical-endogenous-variable multivariate regression are very attractive for quantification and analysis of group judgement data. A controlled feedback approach could be useful in carrying out both an individual-differences scaling and a categorical endogenous-variable multivariate regression, if models tailored to various types of feedback have been developed. That is, a panel of experts might evaluate the relative merits of a collection of R&D projects, perhaps both completed and projected. In the case of individual-differences scaling, group opinions can be fed back, in a multi-stage approach, to produce a scaled solution at each stage. Because the current state of the methodology is strictly mathematical and has not yet advanced to the point where statistical inferences can be drawn, it is difficult to make comparisons of the results at each stage.

In categorical-endogenous-variable multivariate regression, there are advantages in being able to relate the experts' responses to their backgrounds, their degrees of expertise, and their definitions of the "difficulty" parameters of each project, and statistical inferences and predictions are possible. It is recommended that the approach be used without feedback unless the dependency introduced by the feedback is accounted for in the model.[1] Nevertheless, the methodology may be used to advantage to analyse the first-stage results of a controlled feedback process.

In conclusion, methods should be refined for selecting a panel of *appropriate* experts whose opinions will be pooled to form a group judgement. Criteria should be established for determining the number of experts. If a controlled feedback type of analysis is planned, it should first be pretested on the panel of experts. An interrogation procedure should be developed with the questions designed to assess knowledge in specific areas (such as assessment of technological advancement in the specific technologies). Computer programs should be used for analysing data by means of both individual-differences scaling and categorical-endogenous-variable multivariate regression. At least two of the three analytical approaches might be undertaken simultaneously (for supportive validation): individual-differences scaling using a multi-stage controlled feedback approach, categorical-endogenous-variable multivariate regression without the feedback data.

---

1. By using a multi-stage procedure that tells each subject at each stage what the group opinions were at the previous stage, we are in effect generating intrinsic collusion among the subjects. The effect of such collusion upon the statistical analysis is to violate the assumption of formulation previously described will not permit correlated observations independence among the subjects' responses. At the present time the model (interdependent response vectors). Alternative models which compensate for the feedback effects are currently under study and will be reported on at a later date.

Bibliography

Alexander, A.J. and J.R. Nelson, Measuring Technological Change: Aircraft Turbine Engines, R-1017-ARPA/PR, The Rand Corporation, May 1972.

Ament, R.H., "Comparison of Delphi Forecasting Studies in 1964 and 1969," Futures, Vol. 2, March 1970.

American Psychological Association, Standards for Education and Psychological Tests and Manuals, Washington, D.C. 1966.

Bock, R.D., and L.V. Jones, The Measurement and Prediction of Judgement and Choice, San Francisco, Holden Day, 1968.

Brown, B., S.W. Cochran, and N.C. Dalkey, The Delphi Method II: Structure of Experiments, RM-5957-PR, The Rand Corporation, June 1969.

Carroll, J.D., and J.J. Chang, "Analysis of Individual Differences in Multidimensional Scaling via an N-way Generalisations of 'Eckart-Young' Decomposition," Psychometrika, Vol 35, No. 3, September 1970, pp 283-319.

Coombs, C.H., A Theory of Data, New York: John Wiley & Sons, Inc., 1964.

Coppleson, L.W., R.M. Factor, S.B. Strum, P.W. Graff, and H. Rappaport, "Observer Disagreement in the Classification and Histology of Hodgkin's Disease," Journal of the National Cancer Institute, Vol. 45, No. 4, October 1970.

Dalkey, N., and O. Helmer, The Use of Experts for the Estimation of Bombing Requirements, RM-7272, The Rand Corporation, November 1951.

Dalkey, N., and O. Helmer, An Experimental Application of the Delphi Method to the Use of Experts, RM-727 (Abridged), The Rand Corporation, July 1962.

Dalkey, N.C., The Delphi Method: An Experimental Study of Group Opinion, RM-5888-PR, The Rand Corporation, June 1969.

Dalkey, N.C., and B. Brown, and S.W. Cochran, The Delphi Method, III: Use of Self-Ratings to Improve Group Estimates, RM-6115-PR, The Rand Corporation, November 1969.

Dalkey, N.C., B. Brown, and S.W. Cochran, The Delphi Method, IV: Effect of Percentile Feedback and Feedin of Relevant Facts, RM-6118-PR, The Rand Corporation, March, 1970.

Dalkey, N.C., and D.L. Rourke, Experimental Assessment of Delphi Procedures with Group Value Judgements, R-612-ARPA, The Rand Corporation, February, 1971.

Dalkey, N.C., and B. Brown, Comparison of Group Judgement Techniques with Short-Range Predictions and Almanac Questions, R-678-ARPA, The Rand Corporation, May 1971.

Edwards, W., "The Theory of Decision Making," Psychology Bulletin, Vol. 51, No. 4, 1954a, 380-417.

Edwards, W., "The Reliability of Probability Preferences," American Journal of Psychology, Vol. 67, 1954b, 68-95.

Einhorn, H.J., "Expert Measurement and Mechanical Combination," Organisational Behaviour and Human Performance, March 1972.

Guilford, J.P., Psychometric Methods, 2d ed., New York: McGraw-Hill, 1954.

Harman, A.J., assisted by S. Henrichsen, A Methodology for Cost Factor Comparison and Prediction, RM-6269-ARPA, The Rand Corporation, August 1970.

Harman, A.J., Choice Among Strategies for System Acquisition, P-4794, The Rand Corporation, March 1972.

Klein, B.H., "The Decision Making Problem in Development," in R.R. Nelson (ed), The Rate and Direction of Inventive Activity: Economic and Social Factors, Princeton: Princeton University Press, 1962.

Kyburg, H.E., Jr., and H.E. Smokler, Studies in Subjective Probability, New York: John Wiley and Sons, Inc., 1964

Lindley, D.V., Introduction to Probability and Statistics, Vols. I and II, Cambridge: The University Press, 1965.

Marshall, A.W., and W.H. Meckling, Predictability of the Costs, Time, and Success of Development, P-1821, The Rand Corporation, December 1959; also published in R.R. Nelson (ed), The Rate and Direction of Inventive Activity: Economic and Social Factors, Princton: Princeton University Press, 1962.

Meehl, P.E., Clinical vs. Statistical Prediction, Minneapolis: University of Minesota Press, 1954.

Mosteller, F., and P. Nogee, "An Experimental Measurement of Utility," Journal of Political Economy, Vol. 59, No. 5, 1951, 371-404.

Nelson, J.R., and F.S. Timson, Relating Technology to Acquisition Cost: Aircraft Turbine Engines, R-1288-PR, The Rand Corporation, March 1974.

Nerlove, M., and S. James Press, Univariate and Multivariate Log-Linear and Logistic Models for the Analysis of Qualitative Socio-Economic Data, R-1306-EDA, The Rand Corporation, December 1973.

Oppenheim, A.N., Questionnaire Design and Attitude Measurement, New York: Basic Books, Inc., 1966.

Press, S.J., Applied Multivariate Analysis, New York: Holt, Rinehart and Winston, Inc. 1972.

Press, S.J., and A.J. Harman, Methodology for Subjective Assessment of Technological Advancement, R-1375, The Rand Corporation, April 1975.

Preston, M.G., and P. Baratta, "An Experimental Study of the Auction-Value of an Uncertain Outcome," American Journal of Psychology, Vol. 61, 1948, 183-193.

Sackman, H. Delphi Assessment: Expert Opinion, Forecasting, and Group Process, R-1283-PR, the Rand Corporation, April 1974.

Savage, L.J., "Elicitation of Personal Probabilities and Expectations," Journal of the American Statistical Association, Vol. 66, 1971, 783-801.

Summers, R., Cost Estimates as Predictors of Actual Weapon Costs: A Study of Major Hardware Areticles, RM-3061-PR (Abridged), The Rand Corporation, March 1965.

Timson, F.S., Measurement of Technical Performance in Weapon System Development Programs: Subjective Probability Approach, RM-5207-ARPA, The Rand Corporation, December, 1968.

Torgerson, W.S., Theory and Methods of Scaling, New York: John Wiley & Sons, Inc. 1958.

Winkler, R.L., "The Quantification of Judgement: Some Methodological Suggestions," Journal of the American Statistical Association, Vol. 62, 1967, 1109-1120.

Winkler, and A.H. Murphy, "'Good' Probability Assessors," Journal of Applied Meteorology, Vol. 7, 1968, 751-758.

Winkler, R.L., "Probabilistic Prediction: Some Experimental Results," Journal of the American Statistical Association, Vol. 66, 1971, 675-685.

Wirt, J., A. Lieberman, and R. Levien, R&D Management, Methods Used by Federal Agencies, R-1156-HEW, The Rand Corporation, January 1974.

JEAN-PIERRE AUBIN
BERTIL NASLUND

AN EXTERIOR BRANCHING ALGORITHM

Very often a manager is faced with decisions involving many goals such as pollution, cost, risk, etc. He is then looking for a solution which gives him the highest satisfaction possible regarding all of the goals.

The method developed here first computes one reasonably good solution. This is shown to the manager who is asked if he likes the solution or if there is one goal which might be worsened given that all other goals are either improved or unaltered. Thus the manager might say that pollution can go up given that risk and costs may go down and at least are certain not to go up. Then a new solution is computed and shown to the decision maker who is asked the same question as before.

The method finally finds a solution either because the decision maker wants to stop or because no more alterations of the type described above are possible.

7.1 An Exterior Branching Algorithm
by Jean-Pierre Aubin and Bertil Naslund

The Multiple Criteria Problem

It is usually assumed in decision theory that there exists goals U_i (i=1,2,...,n) and that these goals take on different values depending upon the values chosen for certain action variables X_i (i=1, 2, ...,m).

The decision maker if often assumed to have a preference function F(U) which expresses his utility for various values taken by the vector U. If such a function exists and if it can be expresssed in a mathematical form, the multicriteria problem usually is a problem in non-linear programming. The problem however is that the function F(U) may not exist and it may not be possible to give it a sufficiently precise mathematical form.

It is thus difficult to build the function F(U), but, even if it should be possible to do so, it is not always necessary. For instance one only needs to consider efficient actions in a given situation (by an efficient action we mean actions such that there is no other feasible action which gives higher satisfaction along at least one goal dimension U_i while all others are remaining unaltered).

The problem is however to compute those points on the efficient surface that will in the "best" way aid the decision maker in finding the point on the surface where he would like to be. Since this usually means that one must interact with him in various ways, it is difficult to specify what one shall mean by the "best" way. It involves various ways of giving him information during the "road" to the most preferred point, as well as the time it takes to find the point. Another important aspect is how well the decision maker understands the method he is using.

Various methods have been proposed for defining preferences, while the efficient surface is explored. (For a review of these methods as well as multiple criteria methods in general, see Roy (8)). We shall here propose a method for solving that problem but before we start explaining the method two related techniques will be described and in the concluding section a comparison is made between the methods.

Some Previous Work

As was mentioned, it is difficult to express and represent, mathematically or otherwise, ones preferences. This seems either to involve a tremendous amount of work or to result in a very crude and unprecise function. In many situations the whole preference function is not necessary. It is sufficient only to be able to compare efficient actions, two at a time. Below we shall describe two methods that have been suggested in the literature which we shall use in section 5 for comparison with the method developed in sections 3 and 4.

The Method Proposed by Geoffrion (5)

This method assumes that the decision maker has an implicit preference function.

$$F(U_1, U_2, \ldots, U_n) \qquad (1)$$

over the n objectives $(U_1, U_2, \ldots, U_n)$. The function F is assumed to be concave, increasing and differentiable.

Since the function F is not known, it cannot directly be used in the optimization. It is however assumed that the decision maker can provide the marginal rates of substitution in the following way.

$$\left(1,\ \frac{\frac{\partial F}{\partial U_2}}{\frac{\partial F}{\partial U_1}},\ \frac{\frac{\partial F}{\partial U_3}}{\frac{\partial F}{\partial U_1}},\ \ldots\ \frac{\frac{\partial F}{\partial U_n}}{\frac{\partial F}{\partial U_1}} \right) \qquad (2)$$

The vector (2) is co-linear with the gradient at the solution point in question.

Thus for a feasible solution $(X_1^1, X_2^1, \ldots, X_m^1) = X^1$ the decision maker is assumed to provide the vector (2). If this can be done we have information about the direction of the gradient. Various non linear programming methods can then be used to determine a feasible direction to increase F.

Once the direction in which to move has been established, it remains to decide how far one shall go in that direction. Let the distance be d. The new values determined for the decision variables depend upon d and so do all the objectives $U_i[X(d)]$. We can show how all the goals U_i will change, from their present level U_i^1, when the scalar d is changing. How this can look is shown in figure 1 below.

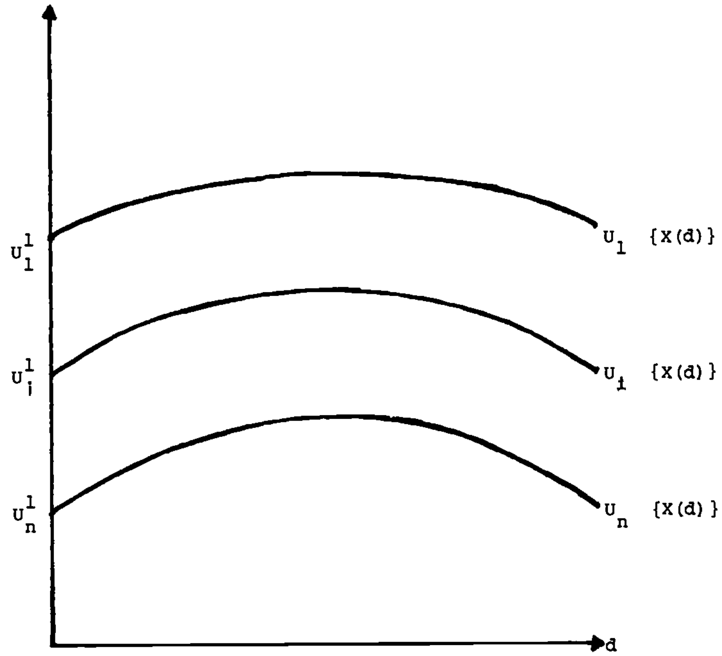


Figure 1

The decision maker is asked to give that value of d which he prefers. This is the second piece of information that he has to provide. Once that value is given a new solution point is obtained and he is again asked to give the marginal rates of substitution etc.

We shall briefly summarise the main steps used in the method (for further details about the use of this method for multiple criteria problems, see Dyer (4)).

(1) Select an initial point X^K. Set $K = 1$

(2) Ask the decision maker about the value of the gradiet at point X^K.

(3) Use that information to find another feasible point, e.g. by using the Frank-Wolfe method.

(4) Use that new point and the point X^K to find a direction d^K in which to alter X^K in order to increase $F(U(X^K))$.

(5) Solve the step size problem

$$\underset{t \geq 0}{\text{Min}} \quad FU(X^K + td^K)$$

by interacting with the decision maker and determine the optimal value t^K. If the new solution $X^{K+1} = X^K + t^K d^K$ is not optimal set $K = K + 1$ and go back to 2.

The Stem Method

The Stem Method developed by Benayoun et. al (2) can only deal with linear relations. Thus the objective functions U_i are of the form.

$$U_i = \sum_{j=1}^{m} a_{ij} X_j \qquad i = 1,2, \ldots, n \qquad (3)$$

and with a feasible region of the form

$$AX \leq b$$

$$X_i \geq o \qquad i = 1,2, \ldots, m \qquad (4)$$

A is a constant n x m matrix

For each objective U_i we determine the optimal value which we denote by M_i. The ideal solution is M which is the point where all objectives take on their optimal values. The values of the decision variables X* which give M are usually not feasible. Therefore a feasible solution X^K is computed. (The index K denotes the k th feasible solution computed). The solution X^K is the one nearest to X* in a minimax sense.

The problem is now to determine

$$\text{Min} \quad \lambda \qquad (5)$$

subject to

$$\lambda \geq \{M_j - U_j (X)\} \quad \Pi_j \qquad j = 1,2, \ldots, n \qquad (6)$$

$$AX \leq b$$

$$X_i \geq o \qquad (7)$$

Thus we see that the difference between the ideal and feasible solution is minimised and that weights Π_j are used to each objective j. The factor Π_j is defined by

$$\Pi_j = \frac{\alpha_j}{\sum_{j=1}^{m} \alpha_j} \qquad \alpha_j = \frac{M_j - m_j}{M_j} \qquad \frac{1}{\sum_{j=1}^{m} a_{ij}}$$

where m_j is the minimum value obtained along goal dimension j when the optimal values for all other goals are determined. $M_j - m_j$ gives an indication of how sensitive the goal is to variations in X and $\frac{1}{\sum_{j=1} a_{ij}}$ normalises the differences (6) for the various goals.

The solution obtained in steps (5) - (7) is shown to the decision maker. If he likes it we are done. If he does not like it, he is asked to select one goal U_K which he is willing to reduce if all other goals are either kept unchanged or increased. He is furthermore asked to specify the amount ΔU_K by which the goal k is reduced. In order to help him determine which k to reduce and the amount of reduction ΔU_K, standard sensitivity analysis in the neighbourhood of the solution to (5) - (7), gives the behaviour of the objective functions U_i (i 1 1,2, ..., n).

Exterior Branching

We make the following assumptions

The problem is

$$\min_{X \varepsilon \chi} \quad F(U(X))$$

U is the n-dimensional vector of real valued functions, X is an m-dimensional vector of real valued variables. The decision-maker is not able to specify F which is his preference function associated with the functions U. χ is assumed to be convex and compact. $U_i(X)$ (i = 1,2,...,n) are convex functions. It then follows that $U(\chi) + \mathbb{R}^+_n$ is a convex subset of $\mathbb{R}^+_n$, (see Aubin (1)). The process which we shall describe does not involve F in any explicit way. We are only going to ask one very simple question from the decision maker at each iteration. The same simple question could in fact be asked from several decision makers, one at each iteration, and the method would still converge to a solution on the efficient surface. As will be discussed in the concluding session, the question about how good the method is has to be judged along several dimensions.

The method can be described by indicating 5 main steps:

1) We determine as in the Stem method the optimal values for each goal dimension disregarding all the others. Thus we obtain an, usually, infeasible solution X^*_j which gives the value α_j to goal j. The point is the shadow minimum.[2] The solution now proceeds using the following steps.

---

1. The method is also described in Aubin (1) and Naslund (6).

2. We consider a minimization problem here while the description of the Stem method was done in a maximisation framework.

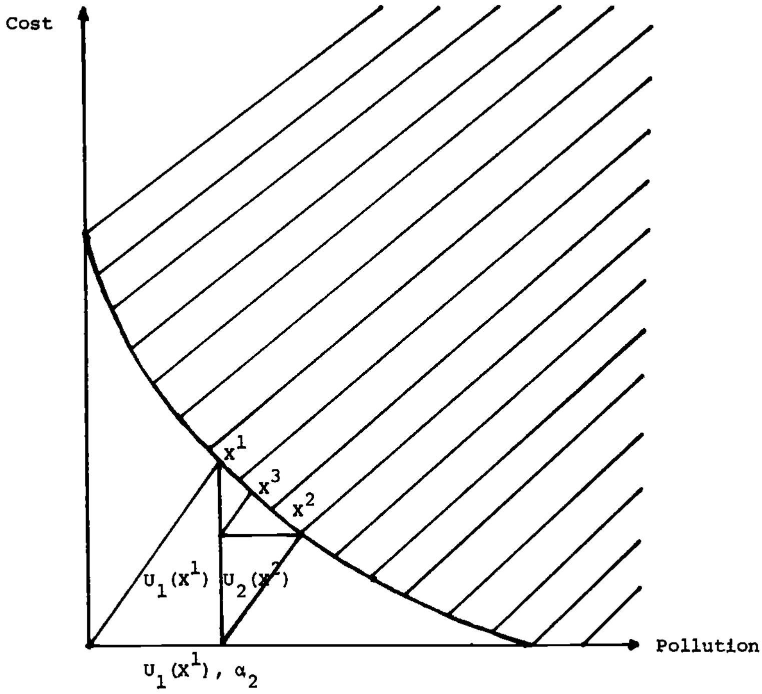


Figure 2

2) The problem

$$\text{Min} \quad \Sigma \, |U_i(X) - \alpha_i|^2 \qquad (8)$$

$$X \in D^1 \qquad (9)$$

is solved. D is the feasible region in the first interaction.
Let the solution be $U_i(X^1)$ ($i = 1,2,\dots,n$). The index 1 indicates the first solution.

3) The solution $U(X^1)$ and the shadow minimum are shown to the decision maker. If he wishes, he can also obtain the marginal rates of substitution at the point X^1. They are using U_1 as the base

$$1, \frac{U_2(X^1) - \alpha_2}{U_1(X^1) - \alpha_1}, \quad \frac{U_3(X^1) - \alpha_3}{U_1(X^1) - \alpha_1}, \quad \dots, \quad \frac{U_n(X^1) - \alpha_n}{U_n(X^1) - \alpha_1} \qquad (10)$$

Thus if the decision maker increases the value of $U_1(X)$ by 1 unit the value of U_2 will be reduced by

$$\frac{U_2(X^1) - \alpha_2}{U_1(X^1) - \alpha_1} \quad \text{etc.}$$

4) The decision maker is asked if he is satisfied with the solution. If he is not, he is asked to mention one goal dimension along which he will reduce his satisfaction provided that his satisfaction along all other goal dimentions will either increase or remain unchanged. In a minimisation problem, we thus ask him which goal he wants to increase given that all others will be reduced or remain unchanged. Assume that he says he can allow U_1 to go up.

5) We next let $U_1(X^1)$ be the new shadow minimum for goal 1, and we solve the problem (compare with (8) and (9)).

$$\text{Min} \sum_{i=2}^{n} |U_i(X) - \alpha_i|^2 + |U_1(X) - U_1(X^1)|^2 \qquad (11)$$

$$X \in D^2 \qquad (12)$$

The new solution $U(X^2)$ is shown to the decision maker and we proceed as in step 2. In section 5 below it is shown that the process converges.

In order to illustrate the method further, let us consider a firm which desires to cause low pollution, U_1, and to operate at low cost, U_2. The decision variables, X, are various production methods and pollution reducing devices. The situation is described in figure 2. The first step provides the solution $U(X^1)$. The decision maker says that he prefers that pollution goes up given that costs go down. We are then able to determine a new shadow minimum $(U_1(X^1)\alpha_2)$ and to compute the new solution $U(X^2)$. The decision maker now thinks that costs can go up if pollution is reduced and we arrive at the new shadow minimum $(U_1(X^1), U_2(X^2))$ and the new solution $U(X^3)$ is computed etc. (see Figure 2).

Precise Statement of the Exterior Branching Method

Let us consider

i) a convex compact set of decisions D

ii) n criteria $U_1(X), \dots, U_n(X)$ (13)

and let

$\alpha^o = (\alpha_1^o, \ldots, \alpha_n^o)$ be the shadow minimum (14)

where

$$\alpha_i^o = \inf_{x \in D} U_i(x) \tag{15}$$

we compute at the first iteration the efficient decision $x^o \varepsilon \chi$ obtained by minimising

$$\sum_{i=1}^{n} |U_i(n) - \alpha_i^o|^2 \text{ over } D \tag{16}$$

Therefore, for computing the second iteration, we choose an index j_1 such that

$$U_{j_1}(x^o) - \alpha_{j_1} > 0 \tag{17}$$

and we define a subset D^1 of the decision set D by

$D^1 = \{x \varepsilon D \text{ such that } U_j(x) \leq U_j(x^o) \text{ for any } j \neq j_1\}$

The new shadow minimum α^1 is defined by

$$\alpha_i^1 = \begin{cases} \alpha_i^o & \text{for any } i \neq j_1 \\ U_{j_1}(x^o) & \text{if } i = j_1 \end{cases} \tag{18}$$

and we obtain $x^2 \varepsilon D^1$ by minimising

$$\sum_{i=1}^{n} |U_i(x) - \alpha_i^1|^2 \text{ over } D^1 \tag{19}$$

Let us assume now that the first iterations $x^o, \ldots, x^{k-1} \varepsilon D$ are computed. For obtaining x^k, we choose an index j_k such that

$$U_{j_k}(x^{k-1}) - \alpha_{j_k}^{k-1} > 0$$

and we define the subset D^k of the decision set D by

$$D^k = \{x \varepsilon D^{k-1} \text{ such that } U_j(x) \leq U_j(x^{k-1}) \text{ for any } j \neq j_k\} \tag{20}$$

the new shadow minimum α^k is defined by

$$\alpha_i^k = \begin{cases} \alpha_i^{k-1} & \text{if } i \neq j_k \\ U_{j_k}(x^{k-1}) & \text{if } i = j_k \end{cases} \tag{21}$$

Then we obtain the k^{th} decision $x^k \varepsilon D$ by minimising

$$\sum_{i=1}^{n} |U_i(x) - \alpha_i^k|^2 \text{ over } D^k \tag{22}$$

In (3), A.M. Charles has proved the following result about the convergence of the method.

Theorem

1. Each decision $x^k \varepsilon D$ defined by (22) is an efficient decision.

2. At each iteration, there exists at least one criterion U_i such that $U_i(x^k) > \alpha_i^k$, and thus, the algorithm is well defined.

3. The distance $\| U(x^k) - \alpha^k \|$ between $U(x^k)$ and α^k is strictly decreasing.

Let us also notice that the sequence of shadow minima α^k is increasing since by the second part of the theorem.

$$\alpha_i^k = \begin{cases} \alpha_i^{k-1} \text{ if } i \neq j_k & \alpha_i^{k-1} \\ & \geq \\ U_{j_k}(x^{k-1}) \text{ if } i = j_k & \alpha_{j_k}^{k-1} \end{cases} \Bigg\} \; \alpha^{k-1} \tag{23}$$

and that the sequence of subsets D^k is decreasing. These remarks and the third statement of the theorem show that the exterior branching algorithm is convergent.

Concluding Discussion

The method described in the third and fourth sections differs from the ones described in the second section in various ways. The Geoffrion method requires that the decision maker can give marginal rates of substitution and step sizes. This has not been required here. Instead we can provide the decision maker with the trade-off possibilities, if he wants to have them, and then he can use them when he decides how to alter his solution.

The Stem method can only solve linear problems and the decision maker is asked to tell by how much he wishes to reduce his satisfaction with respect to one goal dimension provided that the satisfaction along the others, either goes up or remains unchanged. The method developed here is a non-linear method and it does not ask any information from the decision maker except which goal dimension along which his satisfaction can be reduced provided that all others go up or remain unchanged.

It has been shown in many applications of mathematical and other methods to real problems that it is difficult to obtain numerical specifications about goals, restrictions, probabilities, etc. from decision makers (see e.g. Naslund, Sellstedt (8)). The method developed here asks the least from the decision maker.

Since the methods discussed here involve a systematic effort to know the preferences better, it is also important to consider the information that the method gives to the decision maker.

The Geoffrion method gives various solution points. The Stem method gives in addition to that the infeasible shadow optimum and can give the effects of parametric variations of the different goal dimensions. This amounts to the same thing as providing the gradient of the efficient surface.

The method developed here gives the same thing as the Stem method but in addition it gives a new shadow optimum at each step which gradually becomes more realistic and finally (if the decision maker does not stop earlier) ends on the efficient surface. Thus the shadow optimum plays a more important role here than in the Stem method.

Since all the methods require active participation from a decision maker, it seems important to experiment with the methods to see which aspects discussed here are the most important ones and also which method is most easy to understand. It is also often necessary to do experiments in order to determine the speed of convergence of the methods.

References

1. Aubin, J.P. "Closest Efficient Decisions to the Shadow Minimum of a Multiobjective Problem", Working paper 72-6 European Institute for Advanced Studies in Management, February 1972.

2. Benayoun, R, de Montglofier, J, Terguy J, and Larichev O. Linear Programming with Multiple Objective Functions: Step Model (STEM), Mathematical Programming, December, 1971.

3. Charles, A.M. Exterior branching algorithm: a proof of convergence. Cahiers de Mathematiques de la Decision n° 7218. Universite de Paris IX-Dauphine.

4. Dyer, J.S. "Interactive Goal Programming", Management Science, September 1972.

5. Geoffrion, A. "Vector Maximal Decomposition Programming", Working Paper n° 164, UCLA, September 1970.

6. Naslund, B. "A Method for Solving Multi-Criteria Problems", Working paper 72-8 European Institute for Advanced Studies in Management, February 1972.

7. Naslund, B. & Sellstedt, B. The Implementation and Use of Models for R&D Planning Research Policy, Vol. 2, n° 1, 1973.

8. Roy, B. "Problems and Methods with Multiple Objective Functions", Mathematical Programming.

STANLEY ZIONTS

INTEGER LINEAR PROGRAMMING WITH MULTIPLE OBJECTIVES

Although it may seem counterintuitive, a method for solving multiple criteria integer linear programming problems is not an obvious extension of methods that solve multiple criteria linear programming problems. The main difficulty is illustrated by means of an example. Then a way of extenting the Zionts-Wallenius algorithm (4) for solving integer problems is given, and two types of algorithems for extending it are briefly presented. An example is presented for one of the types. Computational considerations are also discussed.

7.2 Integer Linear Programming with Multiple Objectives
by Stanley Zionts

1. *Introduction*

In (4) a method was presented for solving multiple criteria linear programming problems. Because integer programming is a more general case of linear programming, it is reasonable to ask if multicriteria integer problems can be solved in the same way. In this paper it is shown that integer programming problems involving several objectives, where a utility function of the objectives is assumed, can generally be solved using the methods available for noninteger problems provided that the utility function is a linear additive function of the objectives. Using this classification of Roy (1) we therefore restrict the type of problem to class 1, that is, aggregation of multiple objective functions in a unique function defining preferences.

The plan of this paper is to indicate why noninteger methods cannot simply be extended to solve multiple criteria integer problems. Then two extensions of the method of (4) for solving integer problems are developed, an example is solved, and some considerations for implementation, etc, are given. In an appendix the method of (4) is briefly overviewed.

2. *Some Considerations for Solving Multiple Criteria Integer Problems*

The problem to be considered is a mixed integer linear programming problem. Let the decision variables be a vector x of appropriate order where some or all of the variables are required to take on integer values. Denote the set of integer variables as J. The constraint set is then

$$Ax = b \qquad (1)$$
$$x \geq 0$$
$$x_j,\ j \in J \text{ integer}$$

where A and b are, respectively, a matrix and vector of appropriate order. In addition we have a matrix of objective functions C where row i of C gives the i^{th} objective u_i. Each objective of u is to be maximised and we may thus write

$$Iu - Cx \leq 0 \qquad (2)$$

The formulation (1),(2) is the most general formulation of the multiple criteria integer programming problem if one grants that any nonlinearities are already represented in the constraints (1) using piecewise linearisations and integer variables as necessary. If we accept that the implicit utility function is a linear function (as was done originally in (4)) of the objectives u, we may therefore say that our objective is to maximise λ u where λ is an unknown vector of appropriate order. That this problem is an ordinary integer programming problem is trivial. It is also trivial that the problem Maximise λ u subject to (1) and (2), if λ were known, could be solved using any method for solving linear integer programming problems. The problem is that λ is not known.

In an earlier paper (4) Wallenius and I developed a method for solving linear programming problems having multiple objectives. That method is briefly summarised in the appendix. The method has been extensively tested and seems to work in practice. A natural extension of that method would appear to be an extension of solving problems involving integer variables:

1) Solve the continuous multiple criteria problem according to the method of (4)
2) Using the multipliers obtained in 1, solve the associated linear integer programming problem.

Unfortunately as the following simple example shows, that extension does not necessarily work.

Given the constraints:

$$x_1 + \frac{1}{3} x_2 \leq 3\tfrac{1}{8}$$

$$\frac{1}{3} x_1 + x_2 \leq 3\tfrac{1}{8}$$

$$x_1, x_2 \geq 0 \text{ and integer}$$

with objectives $u_1 = x_1$, and $u_2 = x_2$ then provided that the true multipliers λ_1 and λ_2 (>0) satisfy the following relationships

$$\lambda_1 > 1/3\lambda_2$$

$$\lambda_1 < 3\lambda_2$$

then the continuous solution $x_1 = 2.34$, $x_2 = 2.34$ is optimal. However, even for this simple problem there are three optimal integer solutions corresponding to the same continuous optimum depending on the true weights:

If $3\lambda_2 > \lambda_1 > 2\lambda_2$ then $x_1 = 3$, $x_2 = 0$ is optimal

If $2\lambda_2 > \lambda_1 > .5\lambda_2$ then $x_1 = x_2 = 2$ is optimal

If $.5\lambda_2 > \lambda_1 < .5\lambda_2$ then $x_1 = 0$, $x_2 = 3$ is optimal

The example could be readily made more complicated, but it serves to show that further precision is required in the specification of the multipliers than to identify those valid at the noninteger optimum.

3. *Adapting the Zionts-Wallenius method for solving integer programming problems*

To further specify the multipliers λ to find the optimal integer solution, it is necessary to ask additional questions. There are numerous ways in which this may be done, and we shall explore two of them. Both of these proposals represent untested procedures.

3.1 *A cutting plane approach*

The first is a dual cutting plane approach. It is a logical extension of any dual cutting plane method with respect to multiple criteria decision making. Let k be a nonnegative integer, a choice variable, which may be sufficiently large to be effectively infinite. Then the procedure is the following:

1) Find the continuous multiple criteria optimum and set i to 0. Use the associated weights to generate a composite objective function.

2) Adjoin a cut, increment i by one unit and optimise.

3) If the solution is integer, go to 4. Otherwise, if i is not equal to k, go to 2. If i is equal to k go to 4.

4) Set i to zero, generate efficient questions (see the appendix for the definition of efficient variables) for the current solution that are consistent with previous responses. Use the decision-maker's responses to further restrict the multipliers λ and generate the associated composite objective function which may or may not be

different from the old function. If the solution is not integer, go to 2. If the solution is integer, it may or may not be optimal with respect to the new composite function. If it is optimal, stop; if not perform some iterations to obtain a new optimum. If the new optimum is integer stop; if not go to step 2.

That this methid is valid follows from the fact that every time an integer solution is found (and so long as k is not infinite, more often), questions are generated and the multipliers may be altered by the procedure. Thus, for an integer solution the optimality is confirmed or denied. If it is confirmed, the optimal solution has been found; if it is denied, further iterations must be taken. In addition the convergence is assured because of three points:

1) No cut can cut off a feasible integer solution;

2) The corner polyhedron of the interfer solutions has a finite number or extreme points, provided that the solution space is closed and bounded.

3) With any one objective the cut method employed assumes that an integer solution is found with a finite number of cuts.

How well this scheme works depends on the power of the cut method employed. Since dual cut methods are not currently used much because they do not work well in practice, it is unlikely that a multiple criteria scheme based on a dual cut will work well.

3.2 *A branch and bound approach*

We therefore turn our attention to branch and bound algorithms. The multiple criteria method can be altered to work in a branch and bound integer framework. To do this we first present a flow chart of a simple branch-and-bound algorithm, (3), p.416 in figure 1. The idea is to solve a sequence of linear programming problems thereby implicitly enumerating all of the possible integer solutions. The best one found is optimal. Because the multipliers associated with the noninteger optimum are not sufficiently constrained, we cannot use this approach directly here. However, we can modify the approach. The essential change in the algorithm is in block f. We present the motivation for the change as a theorem.

Theorem: A solution can be excluded from further consideration (not added to the list) provided the following two conditions hold:

1) The decision maker prefers an integer solution to it, and

2) All efficient tradeoff questions associated with the solution are viewed negatively or with indifference.

Proof: As shown by the decision-maker's preference the known integer solution has a greater objective function value than the solution in question. Further, since no continuous neighbour is preferred to the solution, any further restricted solution will have a lower objective function than the solution in question and therefore the integer solution.

The result of the theorem may appear unnecessarily strong; however we see no apparent way of weakening it. To use the theorem we alter the test block f of figure 1. We test each of the two conditions of the theorem in that block against the best known integer solution, first with reference to previous responses by solving a very small linear programming problem, and if that is not decisive, by asking the decision maker a question. Each time the decision maker responds to some questions, the solution space on the feasible λ's becomes further restricted. If the old λ's do not satisfy the newly generated constraint(s), then a new set of weights consistent with previous responses must be chosen for further calculation. We also employ a similar com-

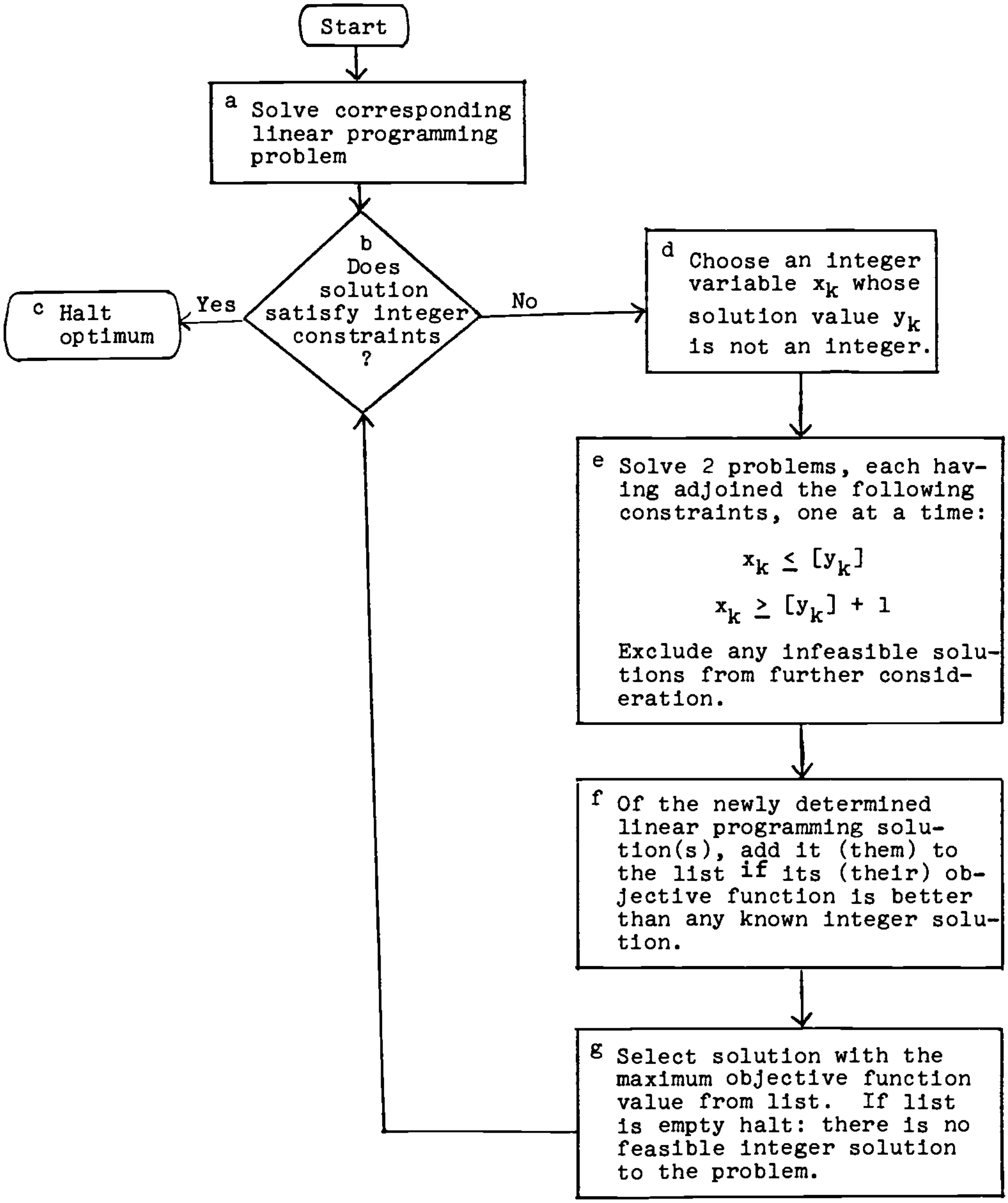


Figure 1

Flow Chart of a Simple Branch and Bound Algorithm

Taken from [5], page 416.

parison in step g between the two solutions found in step e. (This step and some variations are commonly used changes from that given in Figure 1 in branch and bound algorithms). Further, if we have just added two solutions to the list in step e, then we select the solution which is more attractive for futher processing. The solution is so identified first by checking against previous response constraints and if that is not decisive, by asking questions of the decision maker.

The flow chart in Figure 1 is a relatively naive branch and bound model, and other considerations are generally employed as well. For example, as with ordinary branch and bound procedures, finding a good feasible integer solution early is desirable.

We now present an example, the example presented in section 2. We use the naive algorithm with the changes described above assuming that the true weights are $\lambda_1 = .7$, $\lambda_2 = 3$, but that the weights chosen at the continuous optimum are $\lambda_1 = .3, \lambda_2 = .7$. The tree of solutions is given in Figure 2, and the number in each block indicates the order in which each solution is found. (The shaded region is what also would have been generated if every branch had to be terminated either in an integer solution or an infeasible solution without terminating any branches otherwise.)

Tableau 1 is the optimal continuous solution, where x_3 and x_4 are the slack variables. (The identity matrix has been omitted).

Tableau 1

		x_3	x_4
x_1	2.34	1.125	-.375
x_2	2.34	-.375	1.125
u_1	2.34	1.125	-.375
u_2	2.34	-.375	1.125

The questions to Tableau 1 are both efficient (this is not demonstrated) and the two questions are found in the last two rows of the tableau. Are you willing to decrease u_1 by 1.125 units to increase u_2 by .375 units? A simulated response is obtained by using the true weights. Here we compute -1.125(.7) + .375(.3). Since the sum is negative, the response is no. Are you willing to increase u_1 by .375 units by having u_2 decrease by 1.125 units? (Response: no). The negative responses confirm the optimality of the solution of Tableau 1. The constraints are then

$$\lambda_1 > 1/3\lambda_2$$

$$\lambda_1 < 3\lambda_2$$

By using $\lambda_1 + \lambda_2 = 1$, we have, on eliminating λ_2:

$$.25 < \lambda_1 < .75$$

As indicated above we use $\lambda_1 = .3$ (noting that the true value is $\lambda_1 = .7$). Solving the two linear programming problems by branching on x_1 from the noninteger optimum we have solutions 2 and 3. Which is preferred is not obvious and we illustrate the test. Solution 3 has a utility of $3\lambda_1 + .375\lambda_2$. Solution 2 has a utility of $2\lambda_1 + 2.458\lambda_2$. The comparison of the utility of solution 2 less that of solution 2 is

$$\lambda_1 - 2.0833 \overset{?}{\sim} 0$$

On using $\lambda_2 = 1 - \lambda_1$ we have

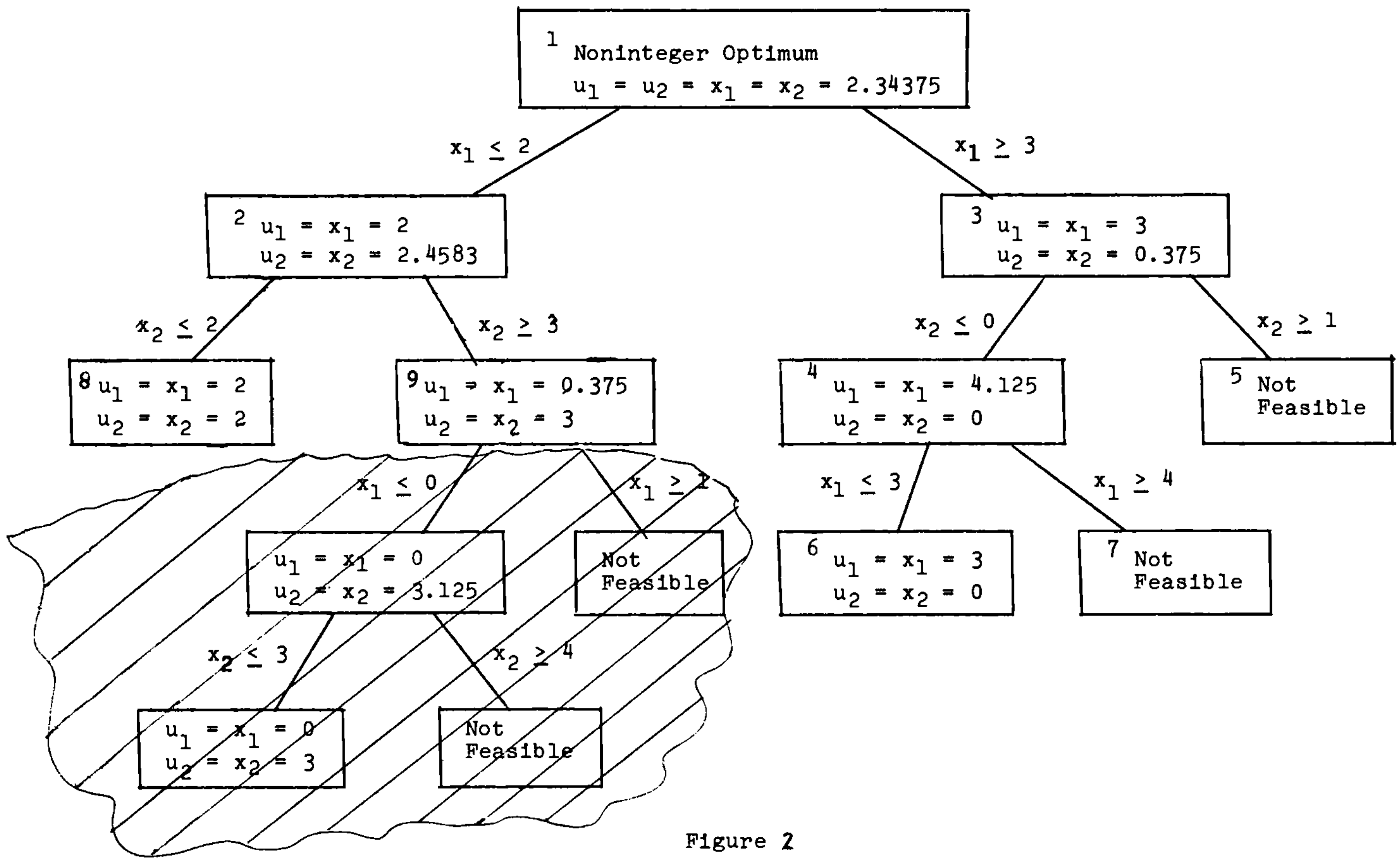


Figure 2

The Branch and Bound Solution

to the Example Problem

$$3.0833\lambda_1 - 2.0833 \overset{?}{\sim} 0$$

Because $.25 < \lambda_1 < .75$, the term can be either positive or negative; hence a question is asked. Since the decision maker prefers solution 3, we have a new constraint.

$$3.0833\ \lambda_1 - 2.0833 < 0 \text{ or}$$

$$\lambda_1 > .675$$

Thus we now have $.675 < \lambda_1 < .75$ so we choose $\lambda_1 = .7$. We then branch on solution 3 to find solutions 4 and 5 (not feasible) and then branch on solution 4 to find solutions 6 and 7 (not feasible). Since solution 6 is integer, we compare it with the only active solution on the list, solution 2. As the answer is not implied, the decision maker is asked which solution he prefers. He prefers solution 6; then the constraint

$$\lambda_1 - 2.4583\lambda_1 < 0 \text{ or } \lambda_1 < .711$$

is added and we have $.675 < \lambda_1 < .711$ and we can continue using $\lambda_1 = .69$. The questions relating to solution 2 have their responses implied to be negative. Hence, solution 2 can be dropped. Since there are no solutions on the list, solution 6 has been found to be optimal. The method of figure 1 using the correct weights enumerates exactly the same solutions.

4. *Discussion*

The implementation of multiple criteria integer programming in liaison with dual cut methods and with branch and bound methods seems straightforward, although it only appears warranted in conjunction with branch and bound methods. It should not be difficult to implement, and it is felt that integer problem would be roughly the same as the performance of a multiobjective linear program as compared to a single objective linear program. More questions will be asked in the integer case, and probably more partial solutions will be generated as well, but it seems that the increase will not be very much. This statement could be made with reference to either of the methods so long as it is not required simply to solve a sequence of integer programming problems as the cut method with k chosen to be large would require. A number of tests which correspond to solving relatively very small linear programming problems must be incorporated, as well. Further statements require testing. For testing, a computer program of the Zionts-Wallenius method now being prepared by the SIDMAR corporation working together with the University of Ghent may be extended to the integer case and used. It is designed to be an easily usable and alterable program.

In the noninteger case we were able to relax the assumption of the linear additive utility function to a general concave utility function. Such a generalisation in the integer case seems unlikely because a point other than an extreme point solution of the corner polyhedron can be optimal. A simple example of such a model would be the use of a utility function involving a product of objectives. (See Bowman (1), for an example). In the linear case a neighborhood would be optimal. Unfortunately, the use of such an idea in the integer case would give an integer solution and a neighbourhood which need not contain any other feasible integer solutions.

APPENDIX: Overview of the Zionts-Wallenius Method (4) for Solving Multiple Criteria Linear Programming Problems

Let the problem of concern be

$$\begin{aligned} Ax &= b \\ x &\geq 0 \\ Iu - Cx &\leq 0 \qquad \text{(A.1)} \\ \text{Maximise } &\lambda u \end{aligned}$$

where $\lambda > 0$ but unknown. The procedure is as follows:

1. Choose an arbitrary set of λ's ($\lambda > 0$).

2. Solve the associated linear programming problem (A.1). The solution is an efficient solution. Identify the adjacent efficient extreme points in the space of the objective functions for which a negative answer by the decision maker is not implied. If there are none, stop; the optimal solution has been found. The marginal rates of changes in the objectives from the point to an adjacent point is a tradeoff offer, and the corresponding question is called an efficient question.

3. Ask the decision maker if he likes or dislikes the tradeoff offerred for each question.

4. Find a set of weights λ consistent with all previous responses. Go to step 2.

REFERENCES

1. Bowman, V.J. "On the Relationship of the Tchebycheff Norm and the Efficient Frontier of Multiple-Criteria Objectives", Paper presented at the conference on Multiple Criteria Decision Making, May 21-23, 1975, at Jouy-en-Josas, France.

2. Roy, B. "Problems and Methods with Multiple Objective Functions", Mathematical Programming, Vol. 1, No. 2, (November, 1971), pp. 239-266.

3. Zionts, S. "Linear and Integer Programming II, Prentice Hall, Englewood Cliffs, New Jersey, 1974.

4. Zionts, S., and Wallenius, J. "An Interactive Programming Method for Solving the Multiple Criteria Problem", Working Paper 74-10, European Institute for Advanced Studies in Management, Brussels, Revised March 1975. Forthcoming in Management Science.

ROBERT E. BROOKS
ASHOK S. KALELKAR

USE OF MULTIDIMENSIONAL UTILITY FUNCTIONS IN HAZARDOUS SHIPMENT DECISIONS

A multi-attribute utility approach is discussed in the context of a case study dealing with the transportation of dangerous chemicals.

8. Use of Multidimensional Utility Functions in Hazardous Shipment Decisions
by Robert E. Brooks and Ashok S. Kalelkar

Introduction

In the last few decades, within the areas of operations research and management science, advances have been made in the development of the theory and application of decision making. This developing sub-field is known as decision analysis.[1,2,3] The aim of decision analysis has been to provide a decision maker in any area of responsibility a logical methodology to evaluate his choices and arrive at the optimum decision. In this paper, the applicability of decision analysis in evaluating choices in hazardous material transportation is evaluated by means of an example involving a hazardous chemical.

The problem of hazardous chemical transportation is a very real one and one that is receiving substantial attention. In recent years both governmental agencies and businesses have made increasing efforts to analyse transportation operations involving chemicals that present potential dangers to people, property, and the environment. Two of the basic questions which must be answered are:

1) Can the proposed transportation operation be made safe enough to be permitted while still meeting other necessary criteria such as profitability?

2) If so, what action (such as choosing a particular mode of transportation) should be taken to minimise the dangers and maximise the benefits provided by the operation?

The simplicity of these questions belies the difficulties encountered in finding the answers. In analysis concerning safety, these difficulties are often associated with measuring risk and quantifying difficult concepts such as the value of life and degrees of adverse impact on the environment. In addition, hazardous chemical transportation sutdies must often deal with hazards that can affect several different sectors of the total environment and for which trade-offs between these sectors and economics must be made.

Part of the first of the questions posed above that asks, "Is the proposed transportation operation safe enough to be permitted?" deals with issues regarding the absolute value of life, property, and the environment. Recently, attempts have been made[4,5] to answer this question, at least in connection with the impact of an operation on human exposures. The general approach has been to evaluate the probability of occurrence of various exposure causing accidents associated with an operation and to examine those probabilities in terms of other risks individuals are willing to take. If individuals are subjected to risk on an involuntary basis or if a single accident has potential of causing several exposures a probability of occurrence several magnitudes lower than once in a hundred years is suggested[4,5] as being an acceptable risk. Very little work has appeared in the area of acceptable environmental risks.

This paper does not concern itself with the first question regarding the issue "how safe is safe enough" or what impacts are acceptable from an individual or government viewpoints. Instead it examines an approach to answering the second question. Assuming that a hazardous chemical operation is "safe enough" to be permitted, which mode of transport should be employed to minimise adverse impacts while maximising the benefits? This paper presents a methodology (by no means unique) for logically optimising the choice of transportation mode based on a particular decision maker's individual preferences and perceptions regarding safety.

General Approach

The methodology of decision analysis presented here is primarily the approach developed by Raiffa[1] and Keeney[2]. The general methodology has been altered slightly to facilitate application to the hazardous chemical transportation problem and to take into account the elements of choice pertinent to the transportation issue in a systematic and consistent way. In general, the steps one takes as follows:

1) Structure the problem by determining what the objectives of the decision maker or decision-making group are (minimise human deaths, environmental damage, and so on).

2) Determine a quantitative performance measure for each objective so that they are operationally defined.

3) Define the set of possible strategies or policies that could conceivably achieve the objectives.

4) From engineering and economic studies determine the range of possible effects that could result from the enactment of the various possible strategies, and as well as possible, determine the likelihood of each of these possible consequences.

5) Assess the decision maker's preferences among all the various consequences, quantify them, and determine which decision will result in the greatest overall acceptability.

This approach can be implemented using a decision diagram in which each potential decision leads to a set of consequences, each of which has a certain probability of occuring and a certain utility to the decision maker. In the case of safety the utilities can be taken to be negative when they refer to injuries to people, property, or the environment. Within these utilities the decision maker reflects his own attitude toward risk and his own judgement of the trade-offs among the various sectors affected.

It may be, for example, that the decision maker is averse to risk and would be willing to pay (through insurance, for example) an amount greater than his long-run expected loss in order to avoid an immediate loss. He may also feel that losses of one type (say human exposures) are generally worse than those of another (say property loss). To quantify these preferences and judgements, the decision analyst can use the questioning technique recently developed for this purpose. This assessment technique, which is presented in an example assessment in this paper, is actually quite simple to apply for most problems and takes only a few hours to complete. Once the decision maker's utility functions have been determined, the worth of each decision can be computed as the sum of the utilities of each possible consequence weighted by the probability that it will happen.

The decision with the greatest numerical value is then the one which best satisfies the decision maker's objectives and hence is his optimal decision.

Basically, this method utilises the experience, judgement and knowledge of the decision maker to aid him in arriving at a decision that best meets his objectives.

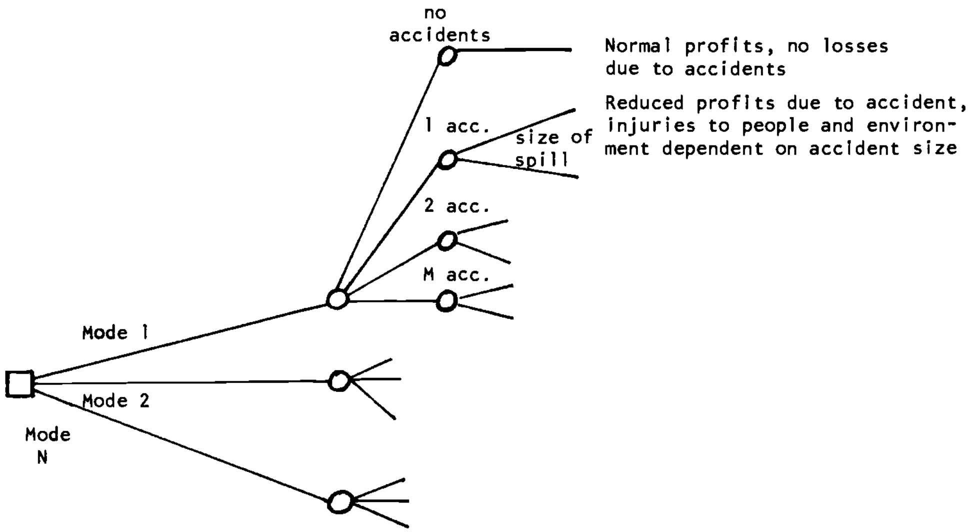
Typical Situation Facing a Decision Maker in Choice of Hazardous Chemical Transportation Mode

A decision maker within a chemical company must decide which transportation mode should be used in shipping a given quantity of a hazardous chemical to a given destination over a definite period of time. Associated with each mode is a certain unit shipment size, unit cost of transportation, and probability of accidental discharge into the surrounding environment. This environment may include persons, property, and/or biological areas, all of which could be harmed by a hazardous chemical spill. The decision maker would like to have a rational means to determine the best mode for his chemical shipments when both costs and risks have been taken into account.

Theoretical Formulation

Decision Trees

The situation described above can be represented by the following decision-tree:



In this formulation a decision maker can choose one of N alternative transportation modes for a period of time T. During this period of time there may be a number of spills. If the frequencey of spills (independent of size) for mode i is given by f_i and if these spills are assumed to occur with exponential interarrival times, then the probability of having n spills in time T is given by the Poisson distribution:

$$P_r\{n;T,i\} = \frac{(f_i T)^n}{n!} e^{-f_i T}$$

Thus the probability of no spill occurring in time T is given by

$$P_r\{0;T,i\} = e^{-f_i T}$$

which when f_i is small compared to 1/T can be approximated as

$$P_r\{0;T,i\} \simeq 1 - f_i T$$

Spill Size Distribution:

If an accident occurs, then some quantity of material will be spilled. Given a set of data representing historical spills one might estimate the probability distribution function g(x) as a log-normal density of the form

$$g(x) = \frac{1}{\sigma x\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(\ln x - \ln x_0)^2}$$

where x_0 is the mode of the distribution, and σ^2 is the variance of $\tilde{x}$. This density has the general form displayed in Figure 1.

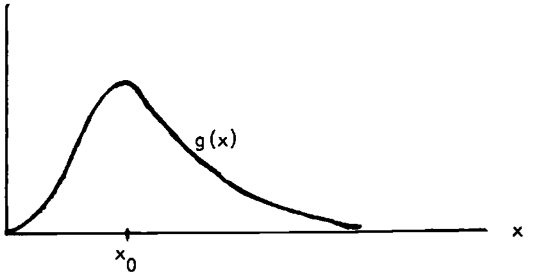


Figure 1: Lognormal Density Function

Given a set of data X_i the parameters x_0 and σ^2 of this density function may be estimated by the unbiased estimators:

$$x_0 = (\prod_{i=1}^{n} x_i)^{1/n}$$

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^{n} (\ln x - \ln x_0)^2$$

Whether this density or another should be used will depend on the actual data in the problem.

Cost/Impact Analysis

In this formulation it is assumed that the demand (Q), the price (P), and the unit costs of manufactures (C_m) and transhipment (C_i) are all known quantities for the chemical of interest. Thus for each mode, i, the gross profits (R_i^0) which accrue to the company, assuming no accidents in the time interval, T, will be given by

$$R_i^0 = (P - C_m - C_i)Q$$

If a single spill of size x occurs in the time T then the gross profits will decrease on two counts: first, at least part of the shipment will be lost; second, damage caused by the spill may have to be repaid. Thus the gross profits could be rewritten as

$$R_i(x) = R_i^0 - Px - D_i(x)$$

Here $D_i(x)$ could be a random variable; under different geographical and weather conditions, for example, the same size spill could cause different amounts of damage. For simplicity, however, we will assume it to be a deterministic function of x.

If several accidents occur the penalties associated with them will be additive since it will be assumed that they will be sufficiently infrequent as to have no interdependent effects.

Thus,

$$R_i(x) = R_i^0 - P \sum_{j=1}^{n} x_j - \sum_{j=1}^{n} D_i(x_j)$$

where x is the n-dimensional vector of spill sizes (x_j).

In addition to these effects on company profits, a spill could have deleterious effects on people, property and/or the biophysical environment in which it occurs. For the moment simply assume that these effects can be quantified in some appropriate units and represented by the two dimensional vectors $\tilde{y} = (\tilde{y}_1, \tilde{y}_2)$, where each $\tilde{y}_i$ is a random variable with a density function dependent on the spill size x. The effect, y, refers specifically to injuries suffered by the external environment. Any monetary costs paid by the company as a penalty for these injuries is included in $D_i(x)$. The accident, however, can be damaging to the company in other ways, its public image, for example. Thus for a particular decision maker within the company the cost/impact value (the disutility) U(x), of an accident of size x will depend in general on two factors: the effect on the company's profits and the external effects. This could be written as:

$$U(x) = h(R,y) = h(R(x),y(x))$$

If more than one accident occurs in the time T, it will be assumed that the associated cost/impact value U(x) can be written as

$$U(x) = U(\sum_{j=1}^{n} x_j)$$

This implies that the decision maker looks at the simple sum of all potential losses over the period T when he evaluates his alternatives. The expected cost/impact value, $ECIV_i$, can then be written as

$$ECIV_i = e^{-f_i T} U_0 + \sum_{n=1}^{\infty} \frac{(f_i T)^n}{n!} e^{-f_i T} U_n$$

where U_0 is the cost/impact value associated with no spill and

where $$U_n = \int_0^\infty \ldots \int_0^\infty dx \ldots dx_n g(x_1) \ldots g(x_n) U(\sum_{j=1}^{n} x_j)$$

Assessment of Cost/Impact Values - An Example

Conser the case where an executive of a small chemical company is faced with transporting 3000 tons of liquid ammonia weekly to a destination 850 miles away. The anticipated operation lifetime is 20 years. For simplicity we assume that the transport tonnage and price remain constant throughout the 20 years (variations in cost can be easily integrated into the decision methodology but are not treated in this example). The decision maker's cost structure is assumed to be as follows:-

Cost of manufacture of liquid ammonia:	\$70/ton
Transportation distance by barge, rail or pipeline:	850 miles
Sale price:	\$150/ton
Cost of transport by barge:	\$9/ton
Cost of transport by rail:	\$15/ton
Cost of transport by pipeline:	\$7/ton

To evaluate the total expected impact value of a hazardous chemical spill, one must first know the range of possible effects. Methods of evaluating quantitatively the extent of hazard are provided by Raju and Kalelkar[6]. In the problem at hand these effects were found to fall within three broad categories: monetary loss due to property damage and lost shipment, environmental damage, and human exposures. A study[7] of the means by which the ammonia could be hazardous (fire in confinement, poisonous vapor cloud, and so on) and an approximate evaluation of the possible consequences of a release provided data from which upper and lower bounds could be determined for the impact magnitudes in each category. Hence appropriate bounds for the assessment of the impact value of a spill were chosen as follows:

1. Monetary loss: $0 - £1,000,000

2. Environment: No effect - death of all plants and animals over a 1 mile stretch of waterway or 100 acres of land

3. People: 0 - 30 exposures

Having placed upper and lower bounds on the possible consequences, we proceed to determine the chosen decision maker's utility or impact value for intermediate values. This function indicates the decision maker's behaviour in the face of uncertain consequences, that is, his attitude toward risk. This attitude can be of three general classes: risk aversion, risk neutrality, and risk proneness. Figure 2 shows these three cases graphically by plotting the utility function (in arbitrary units) against the range of consequences discussed above for each class of impact.

One way to interpret these behaviours follows: A risk-averse person is willing to take a relatively large loss for certain in order to avoid the chance (risk) of an even larger one. For example, a risk-averse person would be willing to pay more than $1000 to avoid a 10% chance of losing $10,000. A risk-neutral person would simply look at expected loss and would be willing to pay exactly that amount (and no more) to avoid an uncertain loss. A risk-prone person would take the chance of losing the $10,000 to avoid the certain $1000 payment, because there is some chance that he will not have to pay anything.

One assesses such a curve by asking the decision maker questions that lead him to find out at what <u>certain</u> value of a given consequence he would not be able to decide between it and a 50/50 <u>gamble</u> between the best and the worst consequence. That is, for all lesser certain impacts he would take the <u>certain impact</u> and for all greater certain impacts he would take the chance on the <u>lottery</u>. Such an impact would be called the <u>certainty equivalent</u> of the gamble lottery.

The first category assessed was monetary loss. In order to make reasonable judgements here, the financial assets of the company have to be taken into account. In this assessment we chose the maximum potential profitability of the ammonia operation with no accidents as the best consequence and bankruptcy of the company as the worst. This range includes all of the financial positions that the company could find itself in with or without accidents. The maximum yearly expected profits for this operation were calculated as $11 million per year. The decision maker's utility function was assessed by asking the following question: if you had to decide between to choices, one of which would give you $10 million for certain and another which would give you a 50/50 chance at $11 million versus zero? Most likely the decision maker would choose the $10 million for certain rather than risking the loss of that $10 million in the gamble. What if the choice were between $10,000 for certain and the same gamble (i.e. $11 million versus $0 at 50/50 odds). In this case he would probably choose the gamble. Thus we know that the gamble is worth more than $10,000 and less than $10 million to the decision maker. By asking the same question with intermediate values one can eventually come to a point where the decision maker is indifferent between a certain payoff of "x" and the 50/50 gamble between $0 and $11 million.

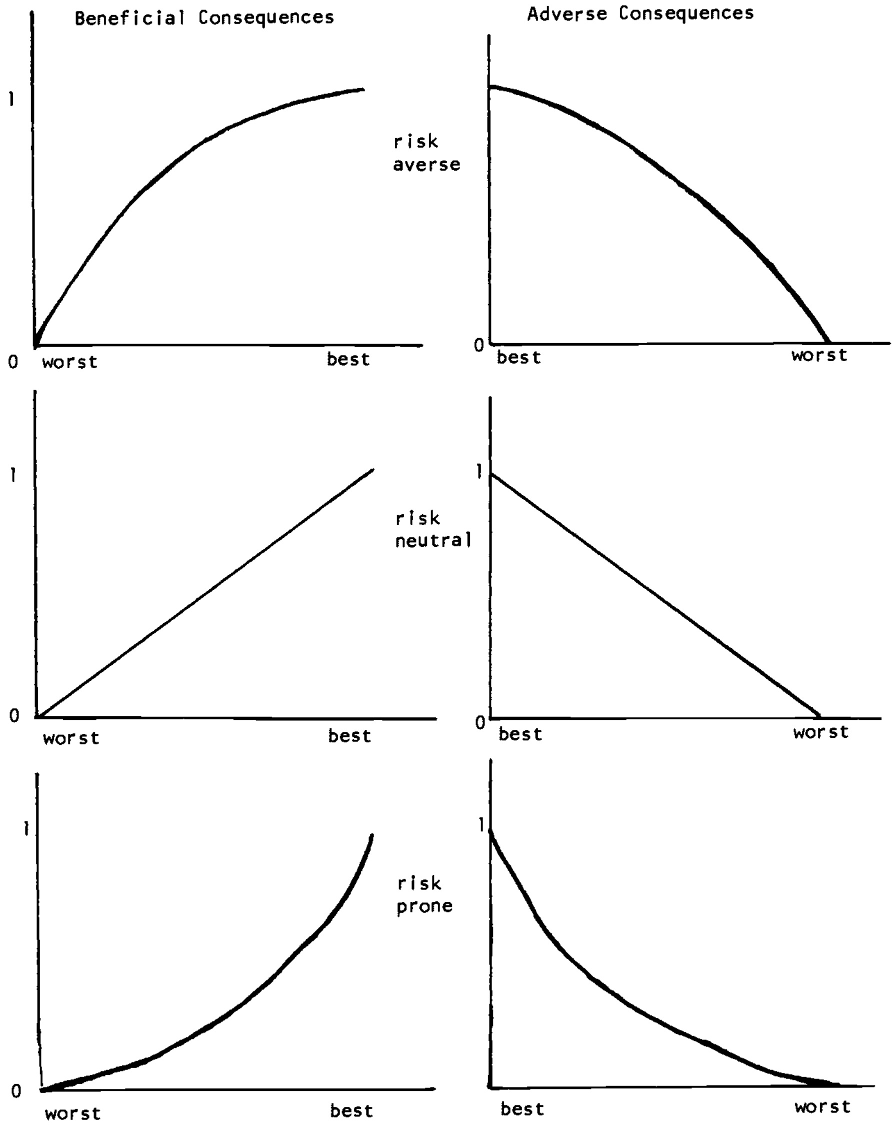


Figure 2: Risk Preference Types

In the example assessment done for this study the decision maker settled on \$1.5 million as his certainty equivalent for the gamble. Since the expected value for the gamble is \$5.5 million the decision maker was seen to be quite risk averse in this choice.

The construction of the decision maker's "utility function" for money can be begun by arbitrarily assigning a value of +1 to \$11 million and 0 to \$0, i.e.,

$$U(\$11) = 1$$

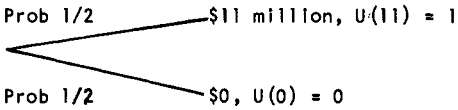
$$U(0) = 0$$

and letting the value of a lottery be given by weighting the value of its consequences by their probabilities. (See Figure 2).

<u>Certainty</u> <u>Lottery</u>

\$1.5 million

$U(1.5) = ?$



Thus in the case at hand, the value of the utility function at 1.5 can be computed as follows

$$U(1.5) = 1/2\ U(11) + 1/2\ U(0) = 1/2$$

A simple curve drawn through the 3 known points of the function U(y) show the decision maker to be risk averse. (Figure 3).

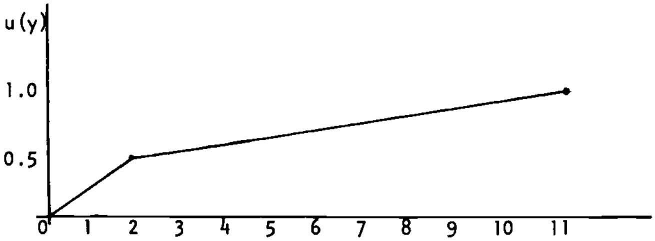
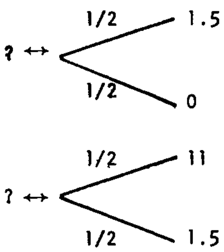


Figure 3

One then determines more points on the curve by assessing the decision makers certainty equivalents for lotteries using the newly determined point (1.5,0.5). That is:



where ↔ means "is valued the same as the gamble _______."

When asked above his preferences for the 50/50 lottery between \$1.5 and 11 million

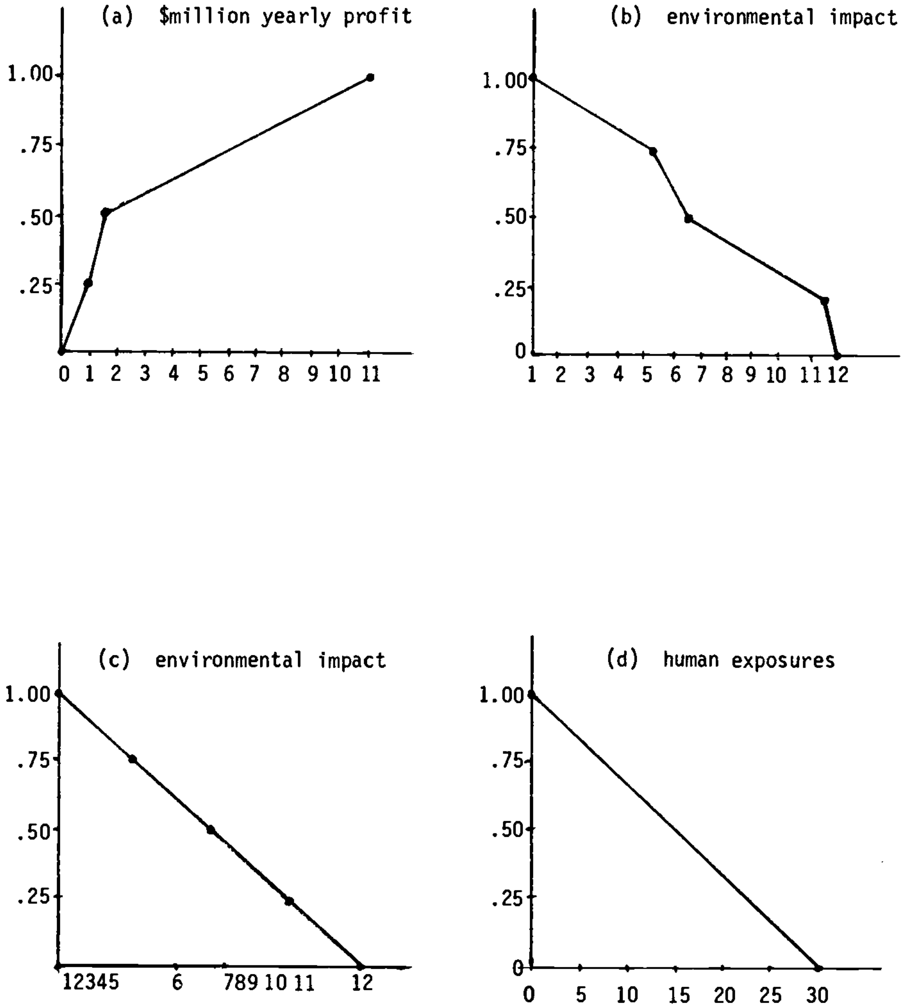


Figure 4: Results of Assessments of Utility Functions

the decision maker said he would "play the odds" i.e., was risk neutral. When dealing with the lower interval between \$1.5 million and bankruptcy, however, the decision maker's certainty equivalent was \$1.0; i.e., as can be seen in Figure 4a, he became risk prone in this region. Why did this occur? As the decision maker explained, the \$1.5 million would be the point at which the value of the company's stock would be worth it to him to take the risk of bankruptcy in order to have the possibility of normal profits and continuing stockholder support. If however expected profits (ie, certain profits) were higher than \$1.5 million, stockholders were getting a greater value for their stock than the plant's worth and therefore he was not willing to risk bankruptcy to get even higher profitability. When the choice was between the \$1.5 million and any higher profits then the decision maker would simply play the odds.

The area of environmental impact has generally been more difficult to assess than money. The approach taken here was to generate first a list of categories of increasingly severe impacts (see Table 1). The more extensive the list, the better and easier it will be to use. One then asks the decision maker to look at the table to decide if the list is monotonically increasing in severity according to his judgement. In addition, if he wishes to subdivide the list even further, that is also allowed. The list and numbering are then adjusted accordingly. At this point the decision maker is asked to find the point on the list such that all categories above this point would be preferred to a 50/50 lottery between no effect and the worst effect, death to all plants and animals over a 100 acre land area or 1 mile of waterway. In this assessment the decision maker chose a point between 6 and 7 as the certainty equivalent. Proceeding as before, the points in Figure 4b were determined.

Note that a simple smooth curve cannot be drawn through the points, if there is a constant interval between categories as there is between exposures (a single exposure) and property damage (a dollar). We have solved this "problem" by drawing a straight line between 1 and 12 and graphing the categories on the horizontal axis using the points generated in the assessment. This procedure yields a measure of the separations between categories. For example, one sees that there are large gaps between 5 and 6, 6 and 7 and 11 and 12 and small ones between 7, 8, 9 and 10. Upon examining the list of categories these results are seen to be quite reasonable. This decision maker cared a great deal more about animal death than plant deaths or injuries. The importantance of having a large number of categories with somewhat "even" intervals is also quite apparent.

The final assessment involved human exposures to ammonia released in a potential accident. In this particular assessment the decision maker stated that he felt quite uncomfortable with finding a certainty equivalent for the lottery of 30 exposures or 0 exposures with 50/50 odds, because he had never been in a position where he had to choose between some number of people exposed <u>for certain</u> and a gamble involving uncertainties. All his choices involved uncertain situations. As will be discussed in the conclusions section this is a weak point in the methodology which needs to be improved.

The decision maker in this example then said that if he had to choose he would "play the odds" i.e., be risk neutral toward this category of effect. Thus his utility or in this case <u>impact value</u> for human exposures would be the linear function graphed in Figure 4d.

Once the assessment of individual impacts has been made, the decision maker is asked to think about the three categories at once in order to determine relative importances and trade-offs between them. In general, this would ve a very difficult task but under certain often valid circumstances Keeney[2] has shown how the numbers can be arrived at in a relatively simple assessment. The requisites for these results are that the different categories of impact (people, profits, and environment) be utility independent in the decision maker's view. This means that utility or impact value curves assessed previously would not be any different if impacts in other areas were different. For example, if the impact value <u>for exposures</u> were assessed knowing that there would be no environmental effect and then again under the assumption of

TABLE 1

ENVIRONMENTAL EFFECTS FROM HAZARDOUS
CHEMICAL SPILLS

1. No effect

2. Residual surface accumlation of harmless material such as sugar or grain

3. Aesthetic pollution (odor-vapors)

4. Residual surface accumulation of removable material such as oil (requires more costly measures of abatement)

5. Persistent leaf damage (spotting, discoloration) but foliage remains edible for wildlife

6. Persistent lead damage (loss of foliage) but new growth in following year

7. Foliage remains poisonous to animals (indirect cause of some deaths upon ingestion)

8. Animals become more susceptiable to preditors because of direct exposure to chemicals and a resulting physical debilitation

9. Death to most smaller animals (consumers)

10. Short term (one season) loss of producers (foliage) with migration of specific consumers (those who eat the specific producer). Eventual reforestation. Death to small animals.

11. Death to producer (vegetation) and migration of consumer (animals). Death to small animals.

12. Death to consumers and producers.

death to all plants and animals in the spill area, the results would be the same. While in some cases this is not true, it was true for the decision maker in this one. Hence Keeney's method and general results apply.

The assessment is done by examining a situation in which one had the best possible consequence on people, profits or the environment, and the worst consequence in the other two and comparing that with a gamble between the best consequence in all three areas and the worst in all three; that is

A: (no exposures, bankruptcy, no environmental effect) for certain versus

B: with probability p → (no exposures, \$11 million profits, no environmental effect)
with probability 1-p → (30 exposures, bankruptcy, death to all consumers and producers

where p is the probability of the best consequence.

The object now is to determine at what probability p* the decision maker becomes indifferent between the two, that is, at $p < p^*$ he would take the certain outcome A and at $p > p^*$ he would take the gamble B.

Keeney has shown that, using the results of assessments of this type, one can compute the total utility (impact value) function as a function of the three (or in general n) variables, which takes on the following simple form:

$$U(x,y,z) = a_1 U_x(x) + a_2 U_y(y) + a_3 U_z(Z) + b_{12} U_x(x) U_y(y) + b_{13} U_x(x) U_z(z) + b_{23} U_y(y) U_z(z) + c_{23} U_x(x) U_y(y) U_z(z)$$

where the coefficients a_i, b_{ij}, and c_{ijk} are determined completely by the numbers p_i^* found in the above assessments and where the $U_i(.)$ are the simple one-dimensional cost/impact functions in Figure 4.

When making these assessments the decision maker made it clear that exposures were the most important category and environment mattered not at all relative to exposures and profits. The function assessed turned out to be

$$U(x,y,z) = 0.9999 U_x(x) + 0.0001 U_y(y)$$

where x is the number of exposures and y is profits.

All cross product and environmental terms dropped out of the function. The coefficients in the above equation are related to the trade-off assessed between exposures and profits, in which the decision maker said that in the choice between a certain 30 exposures with \$11 million profits and a gamble with probability p* of no exposures and \$11 million profits and probability 1-p* of 30 exposures and bankruptcy, he would take the gamble as long as there was 1 chance in 10,000 that the 30 exposures could be avoided, even if that meant a near certainty of bankruptcy.

Modal Analysis

Now that we have a cost/impact function for the three attributes we can look at the frequencies and distributions of accidents and the resulting consequences and compute an expected cost/impact value for each mode. These will then serve as a basis on which to compare the overall desirability of each mode.

Reviewing Figures 4a and 4b we may write the utility functions U_x and U_y as

$$U_x(x) = 1 - x/30 = 1 - 0.03333x, \; 0 \leq x \leq 30$$

$$U_y(y) = 1/2 + (y - 3/2)/19 = 0.42105 + 0.05263y \; 1.5 \leq y \leq 11$$

Since in all the following calculations both x and y are found in these ranges, the other portions of the curves will not be needed.

For the problem at hand it has been determined that accidents occur as (0,1) events; that is, either an accident results in no spill or a complete spill. Thus we can consider the frequency of spills to be the frequency of accidents times the fraction of accidents in which there is a spill. In addition the density of spill sizes is simply the delta function* $\delta(a - a_0)$. That is, a_0 is the only possible spill size for mode i. In each of the following cases an estimate of the expected value of the distribution of exposures and monetary loss due to a spill has been made. Since the functions U_x and U_y are linear over the entire range of these distributions, only the meaning the distribution is needed to determine the expected utility. This is because for any density f(x) it is true that

$$\int_{\infty}^{\infty} f(x) \left[mx - c\right] dx = m\bar{x} - c$$

where $\bar{x}$ is the mean of the distribution. Thus in the case of utility functions which are risk neutral over the range of interest, the exact form of the densities is not needed, only the mean.

Pipeline Analysis

The parameters assumed for the pipeline in this example are as follows:

Transportation cost: \$7/ton
Maximum profit: \$10.950 million**
Spill size: 80,000 gallons (220 tons)
Frequency of spill: 1 every 4.3 years
Mean number of exposures: 5 (rural)
Mean monetary loss: \$10,000 (damage) + \$33,000 (value of ammonia at \$150/ton) = \$43,000 (per accident)
Environmental: type 10 consequence over 15 acres

The pipeline is considered to be located entirely within rural areas. Since the frequency of spill is not very low we will compute the cost/impact value for 1, 2, and 3 accidents in a one year period.

One Accident

In this case $\bar{x}$, the average number of exposures, is 5 and the average profits, $\bar{y}$, is \$10,907 million. Thus the cost/impact value, CIV, for a single accident is

---

*A delta function has the property of being 0 at every point other than at 0 where it is infinite. In addition

$$\int_{a}^{a} \delta(x) f(x) dx = f(0) \quad a > 0$$

It can be derived from the log normal distribution with parameters a_0 and σ^2 by letting $\sigma^2 \to 0$.

**Manufacturing cost is \$70/ton. Thus profit per ton is \$150 - 70 -7 = \$73. Yearly tonnage is 150,000; thus profit is \$10.95 million/year.

$$CIV_1 = 0.9999\ U_x(\bar{x}) + 0.0001\ U_y\ (\bar{y})$$

$$= 0.9999\ (1 - 0.0333 \times 5) + 0.0001\ (0.42105 + 0.05263 \times 10.907)$$

$$CIV_1 = 0.83335$$

Two Accidents

In this case it can be shown that $\bar{x} = 10$ and average profits $\bar{y} = 10.864$. Thus

$$CIV_2 = 0.66670$$

Three Accidents

Here $\bar{x} = 15$, $\bar{y} = 10.821$

Therefore,

$$CIV_3 = 0.50005$$

The frequency of pipeline accidents is 1/4.3 = 0.23 per year. The probability of 0 - 3 accidents in a year is given by

$$P(0) = \exp(-0.23) = 0.7925$$

$$P(1) = 0.23 \exp(-0.23) = 0.1843$$

$$P(2) = 1/2\ (0.23)^2 \exp(-0.23) = 0.0214$$

$$P(3) = 1/6\ (0.23)^3 \exp(-0.23) = 0.0017$$

$$P(4+) < 0.0001$$

The expected cost/impact value ECIV for pipeline is thus given by

$$\text{ECIV (pipeline)} = P(0) \times 1^* + P(1) \times CIV_1 + P(2) \times CIV_2 + P(3) \times CIV_3$$

$$= 0.96120$$

Barge Analysis

The barge is assumed to spend approximately 15% of its time in urban waterways and 85% in rural areas. The parameters associated with barges are assumed to be:

Transportation costs: $9/ton
Maximum profit: $10.650 million
Spill size: 3000 tons
Frequency of spill: 1 every 560 years
Mean number of exposures: 15 (urban); 5 (rural)
Mean monetary loss: $700,000 (urban); $475,000 (rural)
Environmental: Type II over 1 mile of waterway

Because barge spills are so infrequent we need only compute the result for a single spill. We will do this for urban and rural separately and then combine them according to the relative probability of finding the barge in the respective locations.

---

* Zero accidents has a cost/impact value of 1.0 since there will be no exposures and maximum profit.

Urban

In this case $\bar{x} = 15$, $\bar{y} = 9.950$

Thus, $CIV_{urb} = 0.9999\ (1 - 0.0333 \times 15) + 0.0001\ (0.42105 + 0.05263 \times 9.95)$

$= 0.50004$

Rural

Here $\bar{x} = 5$, $\bar{y} = 10.175$

$CIV_{rur} = 0.9999\ (1 - 0.0333 \times 5) + 0.0001\ (0.42105 + 0.05263 \times 10.175)$

$= 0.83335$

Averaging these over the entire journey gives

$$CIV = (0.15)(0.50004) + (0.85)(0.83335)$$

$$= 0.78335$$

The frequency of barge spills is given by 1/560 = 0.0018.

Thus

$$P(0) = \exp\ (-.0018) = 0.99822$$

$$P(1) = .0018 \exp\ (-.0018) = 0.00178$$

$$P(2+) < 0.000002$$

The expected cost/impact value of the barge mode is therefore

$$ECIV(\text{barge}) = P(0) \times 1 + P(1) \times CIV$$

$$= 0.99961$$

Rail Analysis

In the analysis of the rail mode, the fraction of time spent in urban areas is taken as 0.19 and for rural areas 0.81. The parameters are assumed to be:

Transportation cost: $15/ton
Maximum profit: $9.75 million
Spill size: 30,000 gallons (82.5 tons)
Frequency of spill: 1 every 15 years
Mean number of exposures: 5 (urban), 1.5 (rural)
Mean monetary loss: $22,000 (urban or rural)
Environmental: type 10

In this case accidents can occur in both urban and rural areas and the frequency is high enough that two accidents could conceivably occur in one year. Thus the following possibilities will be considered

<u>Urban - 1 accident</u>

$\bar{x} = 5$, $\bar{y} = 9.728$

$CIV_{U,1} = 0.83334$

Urban - 2 accidents

$\bar{x} = 10, \bar{y} = 9.706$

$CIV_{U,2} = 0.66670$

Rural - 1 accident

$\bar{x} = 1.5, \bar{y} = 9.728$

$CIV_{R,1} = 0.95000$

Rural - 2 accidents

$\bar{x} = 3.0, \bar{y} = 9.706$

$CIV_{R,2} = 0.90000$

Urban and Rural Accidents

Here $\bar{x} = (0.19)(5) + (0.81)(1.5) = 2.165$

$\bar{y} = 9.706$

$CIV_{R,U} = 0.92783$

The frequency of accidents is 1/15 = 0.067 per year. Thus

$P(0) = 0.93551$

$P(1\text{ urban}) = (0.19)(.067)(.93551) = 0.01185$

$P(1\text{ rural}) = (0.81)(.067)(.93551) = 0.05052$

$$P(2\text{ urban}) = (0.19)^2 \frac{(.067)^2}{2} (.93551) = 0.00008$$

$$P(2\text{ rural}) = (0.81)^2 \frac{(.067)^2}{2} (.93551) = 0.00136$$

$$P(\text{urban + rural}) = (.19)^2(.81) \frac{(.067)^2}{2} (.93551) = 0.00032$$

$P(3+) < 0.0004$

Expected cost /impact value

ECIV (rail) = 0.99495

Rankings

Now that expected cost/impact parameters have been computed for each mode, they can be ranked. (See Table 2).

The last column measures the difference in expected cost/impact value from the risk-free value of 1.00000. It indicates that rail is an order of magnitude more risky (to the DM) than barge, while pipeline is two orders of magnitude greater. The recurrence intervals follow approximately the same pattern. The cost/impact value for a single accident indicates that barge is riskier. This is because it carries the largest single shipments. However, the great difference in recurrence intervals more than makes up for the increased risk caused by greater shipment size. In the case of the decision maker utilised for the example problem, barge would be his optimum choice of transport.

TABLE 2: RANKING BY MODE

Rank	Mode	CIV for Single Accident Urban	CIV for Single Accident Rural	Recurrence Interval	ECIV	1-ECIV
1	Barge	.50004	.83335	560 yr	.99961	.00039
2	Rail	.83334	.95000	15 yr	.99459	.00505
3	Pipeline	-	.83335	4.3 yr	.96120	.03880

Discussion

The decision analysis methodology presented here makes use of multi-dimensional utility functions and can be a useful aid to a decision maker in choosing his optimum strategy. It should be realised that results presented in the ammonia example reflect the view of one decision maker and are subject to change with different decision makers. The main purpose of the example was to demonstrate the availability of a quantifiable methodology for use in decision making in the field of hazardous materials transport. In the example presented, all cost and accident figures were assumed for illustrative pruposes and should not be construed as being representative of either the chemical or the transportation mode.

In the process of developing this paper three major issues surfaced which may limit the applicability of this methodology in some situations. These issues are as follows:

1. The utility assessment methodology itself suffers from a very real difficulty. the procedure for developing a decision maker's preference or utility function employs a questioning technique involving certainty equivalents to probable occurrences. Whereas this approach is viable for guaging profit considerations and impact considerations involving property and environment damages it is not realistic for human exposure impact evaluation. This is because decision makers in business do not actually deal with situations involving consequences where there will be a number of human exposures "for sure". Thus the idea of certainty equivalent in this context is not sensible to them. This contrasts with the money dimension where money for sure versus gambles in money are very real indeed. A better assessment technique involving all probabilistic tradeoffs rather than certainty equivalents is needed for cases where a "certain" consequence is an unreal alternative on which to base a decision. This would be especially true if human deaths and injuries were involved.

2. The more detailed information one has on transportation statistics the more valid will be the optimum choice. Accident and spill rates for specific chemicals over different transportation routes and modes are not easily available. Several Federal Government agencies have started collecting accident/spill data but the data base is not always complete. Ideally, a probability distribution of spill sizes is needed to perform a rigorous analysis. In the absence of such data one may resort to experience in related fields, national average data, engineering judgements or just plain guesses. A large uncertainty in the data base will result in a final choice of questionable validity.

3. Finally, in converting spill volumes (on land or water) into extent of damage to people, property and the environment several simplifying assumptions have to be made. Many of the problems associated with damage assessment are discussed in referent 6 and 8. Once again, unless realistic modeling of the

consequences of a spill is performed the end result may suffer as far as validity of choice is concerned.

In conclusion, it is stressed again that the purpose of this paper is to present a quantitative methodology for optimizing choice in the face of several options. The methodology does not address the question of whether a given operation should be permitted in the first place. The application of the methodology to choice of mode of transport in shipment of hazardous chemicals is found to be relatively successful. The problems associated with the treatment of adverse human exposure needs further work in order to tighten the methodology and make it more applicable to the transportation problem.

Acknowledgement

The authors gratefully acknowledge the generous time and valuable insight provided by the corporate vice president of a major chemical company who acted as a decision maker for the example problem.

REFERENCES

1. Raiffa, H., Decision Analysis, Addison Wesley, Reading, Mass. (1968).

2. Keeney, R.L., "Multidimensional Utility Functions: Theory Assessment and Application," TR No. 43, Operations Research Center, MIT, Cambridge, Mass. October 1969.

3. Kalelkar, A.S., Partridge, L.J., and Brooks, R.E., "Decision Analysis in Hazardous Material Transportation," Proceedings of the 1974 National Conference on Control of Hazardous Material Spills. (Pages 336-344) American Institute of Chemical Engineers, San Francisco, California, August 1974.

4. Starr, C., "Social Benefit versus Technological Risk," Science 19 September 1969, p. 232.

5. Rasmussen, N., "Reactor Safety Study," Report WASH-1400, Atomic Energey Commission, 1974.

6. Raj, P. and Kalelkar, A., "Models in Support of the Hazard Assessment Handbook," USCG Report CG-D-65-74, (NTIS #AD776617), January 1974.

7. Raj, P., Hogopian, J. and Kalelkar, A., "Prediction of Hazards of Spills of Anhydrous Ammonia on Water," USCG report CG-D-74-74, (NSIS#AD779400), January 1974.

8. "Hazard Assessment Handbook," Manual 3 of Chemical Hazard Response Information Systems, USCG Report CG-446-3, January 1974.

RAYMOND TREMOLIERES

AUTOMATING DECISION MAKING FOR THE DIAGNOSIS PROBLEM

A probabilistic decomposition of the diagnosis problem is used to derive an optimal decision model in the medical context.

9. Automating Decision Making for the Diagnosis Problem
by Raymond Tremolieres

Introduction

In this paper we report some results of research into the automation of the diagnosis problem. This problem was investigated as part of a research effort in computer-aided diagnosis.[1] Herein we give two[2] formulations for the undercriterion diagnosis problem;

- the deterministic diagnosis problem;
- the probabilistic diagnosis problem.

In order to facilitate the understanding of the text, in the foreword we provide an explanation of the diagnosis problem in medical terms. Although the words used are medical ones, the formalism has more general applicability, and perhaps its usefulness in problems of medical decision making will stimulate interest in applying it in other management areas.

The Medical Problem

For simplicity, we consider three diseases, D_1, D_2, D_3, and we assume that these diseases can be detected by several tests or questions, for example, T_1, T_2, T_3, T_4. Each test gives several possible answers, for example[3]

$T_1 = T_1^+$ or T_1^- (+ for "yes", - for "no")

$T_2 = T_2^+$ or T_2^-

. . .

$T_4 = T_4^+$ or T_4^-

The various combinations generated from the results correspond to the disease semeiologic profiles. Thus, we can establish a correspondence table between diseases and the results of the tests.

> The diagnosis process consists of finding the subsets of tests that enable one to distinguish the diseases; these subsets are called filtering. They must fit, as close as possible, the criteria fixed by the physician: costs, performing time, physical disagreement; that is, the so-called undercriterion diagnosis.

The Logical Problem

The main characteristic of the logical problem is that the result of a test is <u>deterministic</u>. In other words, if somebody has the disease D_1, then this corresponds to a very specified set of answers. For example, without any doubt, the answer vector must be $(T_1^-, T_2^+, T_3^+, T_4^+)$ (see Table 1) and if somebody else answers $(T_1^-, T_2^+, T_3^+, T_4^-)$, then he cannot have the disease D_1.

---

1. Work supported by D.G.R.S.T. convention 70 02 190.
2. One is a little extension of the Mattei-Faure-Yacoub's Model (1).
3. In what follows we suppose that there are only two answers for each test: yes (+) or no (-). However, if it is necessary, generalisation of more than two answers would not be difficult.

For simplicity we consider the following correspondence table between tests and diseases (we write "1" for +, "0" for -).

Time	Cost	Questions	Diseases		
			D_1	D_2	D_3
2	9	T_1	0	1	1
5	1	T_2	1	0	0
3	4	T_3	1	1	0
1	5	T_4	1	1	0

Table 1: In case (T_1, D_1), 0 = - = no to question T_1; in case (T_1, D_2), 1 = + = yes to question T_1.

We say that a subset of the test set (T_1, T_2, T_3, T_4) is a <u>filtering</u> set of tests, or that the tests considered are <u>discriminating</u>: if they are sufficient to say that the patient, if he has one of the three diseases, has one well-determined disease and not one of the others.

For example:

(T_1, T_2) are not discriminating because they are not sufficient to distinguish D_2 from D_3.

but

(T_1, T_3) and (T_2, T_3) are discriminating.

If some subset is filtering, then obviously any other subset that contains it is also a filtering set.

When a subset $\mathcal{J}$ is a filtering set and when there does not exist another filtering subset that is contained in $\mathcal{J}$, we say that $\mathcal{J}$ is a <u>basis filtering subset</u>. We call $\mathcal{P}*$ the set of all the basis filtering subsets of (T_1, T_2, T_3, T_4).

The <u>diagnosis problem</u> consists of finding all the basis filtering subsets, i.e., in defining $\mathcal{P}*$.

The undercriterion diagnosis problem consists in finding in $\mathcal{P}*$ all or one of the filtering subsets that optimise a given criterion.

For example we get

$$\mathcal{P}* = \{(T_1, T_3), (T_1, T_4), (T_2, T_4)\}$$

and (1) if the main criterion is the cost, then we find that (T_2, T_4) is the minimal cost basis filtering subset: 1 + 4 = 5

(2) if the main criterion is the time, then the solution is (T_1, T_4): 2 + 1 = 3.

Method to Solve the Logical Problem

First of all we construct a special table called the "and" table. In this table we write for each couple of diseases the tests that enable us to distinguish the two elements of the couple (see Table 2). After this we define the "and ordering" table when ordering the lines of the "and" table by increasing the number of tests. Thus we get Table 3. Now we are going to show how it is possible to compute the filtering subsets by a method that we call the global drawing method. In this method we establish a tree, where the strings of elements which link the tops to the roots are the desired filtering sets. The method is as follows:

First step: At level 1 of the tree, we write the tests of line L_1 (see Table 3). The elements T_1, T_2, put on this first level, are called the roots of the tree (see Figure 1).

Second step: At level 2, we write all the tests of line L_2 under each element of level 1 as many times as there are elements at level 1, but under the condition that we don't write the tests of L_2 under an element of L_1 if this last element is also in L_2. Thus we write T_3, T_4 two times (See Figure 2). As we link by an arc the elements of L_2 to the element of L_1 under which they are written.

At any stop of the drawing of the tree, we call a node a hanging node if this node is not linked to a node situated on a higher level (down in the figure).

Third step: Under each hanging node, we write on the next level all the elements of L_3, except if the string that links the hanging node to the first level contains an element that is also in L_3.

For this reason we don't write L_3 here.

Next steps: For L_4, L_5, L_6,..., we proceed as for L_2, until we have used all the lines of the "and ordering" table.

In the last step the strings that link the terminal nodes to the roots give all the basic filtering subsets (but sometimes other filtering subsets) (see Figure 2 and Table 4). Comparison between the values of the filtering subsets obtained by the global drawing method give the minimal cost or minimal time filtering subsets (see Table 4).

The cost minimising filtering subset is (T_2, T_3) and the time minimising filtering subset is (T_1, T_3).

A filtering subset that approximately minimises the cost and the time is (T_2, T_4) (see Table 4).

Special methods for the undercriterion diagnosis problem are given in Begon-Tremolieres (1).

To distinguish	we may use the tests
↑ D_1 from D_2	T_1 or T_2
and D_1 from D_3	T_1 or T_2 or T_3 or T_4
↓ D_2 from D_3	T_3 or T_4

Table 2: "and" table

To distinguish	we may use the tests	
D_1 from D_2	T_1 or T_2	L_1
and D_2 from D_3	T_3 or T_4	L_2
D_1 from D_3	T_1 or T_2 or T_3 or T_4	L_3

Table 3: "and ordering" table

Level 1 . . . T_1 . . . T_2 . . .

Figure 1

Level 1 . . . . T_1 . . . . T_2 . . .

Level 2 . . . T_3 . . . T_4 . . . T_3 . . . T_4 . .

Figure 2

No.	Filtering subsets given by the global drawing method	Time	Cost
1	T_1 T_3	5	13
2	T_1 T_4	3 *	14
3	T_2 T_3	8	5*
4	T_2 T_4	6	6

Table 4

The Probabilistic Problem

Herein we remove the hypothesis done previously: the result of a test is not deterministic for a disease. Namely we just know, for example, that:

> among the people that have the disease D_1, 0.1 percent usually answer yes to question T_1.

This leads us to establish the probability table between tests and diseases. In this table, the element in case (T_i^+, D_j) is $P(T_i^+/D_j)$ which is the probability to say yes to question T_i if the patient has the disease D_j (see Table 5), the probability to say "no" to the questions T_i being $P(T_i^-/D_j) = 1. - P(T_i^+/D_j)$.

Time	Cost	Answers	Diseases		
			D_1	D_2	D_3
2	9	T_1^+	0.1	0.6	0.90
5	1	T_2^+	0.95	0.0	0.0
3	4	T_3^+	0.98	1.0	0.0
1	5	T_4^+	0.9	0.88	0.2
"a priori" probabilities			0.6	0.1	0.3

Table 5

In what follows we give a simplified approach to the problem.

We suppose given a scalar $\varepsilon > 0$ sufficiently small.

We say that a subset S of (T_1, T_2, T_3, T_4) is a probabilistic filtering subset if, for each D_j, j=1,2,3, there exists a realization of S, that we write S_j, such that

$$P(D_j/S_j) \geq 1 - \varepsilon$$

$$P(D_k/S_j) \leq \varepsilon \text{ , } \forall k \neq j$$

Moreover, we say that S_j is characteristic for the disease D_j.

We say that a set S is a basis probabilistic filtering subset if (1) it is a probabilistic filtering set and (2) there is no probabilistic filtering set strictly contained in S.

The probabilistic diagnosis problem consists of finding all the basis probabilistic filtering subsets. We call P^* the set of all these subsets.

The undercriterion probabilistic diagnosis problem consists in finding in P^* all or one of the probabilistic filtering subsets that optimise a given criterion.

To know if a subset S', for example

$$S' = (T_1^-, T_2^+, T_3^+)$$

is a p-filtering subset (probabilistic filtering subset) we must compute

$P(D_j/S')$, $j = 1,\ldots,3$

This can be done by using the Bayes' formulae:

let S' by a subset of n elements ($n \geq 2$), and

let $S' = (\tilde{S}', T')$ where $\tilde{S}'$ is any subset of S' that contains n-1 elements and T' is the remaining element.

Then the Bayes' formula is

$$P(D_j/S') = \frac{P(D_j/\tilde{S}') \times P(T'/D_j)}{\sum_{k=1}^{3} P(D_k/\tilde{S}') \times P(T'/D_k)}$$

where S' contains only one element, the formula is

$$P(D_j/S') = \frac{P(D_j) \times P(\dot{S}/S_j)}{\sum_{k=1}^{3} P(D_k) \times P(\dot{S}/D_k)}$$

where $P(D_k)$ is the "a priori" probability of the disease D_k (see last line of Table 5).

In Table 6 we give the conditional probabilities of all the possible subsets of results for T_1, T_2, T_3.

It is not difficult to see that there are only two p-filtering subsets (with $\varepsilon = 0.1$), namely $\{T_1, T_2, T_3\}$ and $\{T_1, T_2, T_3, T_4\}$. We give the different results for these two subsets in Tables 7 and 8.

T_1	T_2	T_3	Conditional Probabilitis of D_1	D_2	D_3	
+	+	+	1.	0.	0.	
+	+	-	1.	0.	0.	
+	-	+	0.05	0.95	0.	
+	-	-	0.0002	0.	0.9998	
-	+	+	1.	0.	0.	
-	+	-	1.	0.	0.	
-	-	+	0.40	0.60	0.	*
-	-	-	0.02	0.	0.98	

Table 7

Conditional Probabilities

$T_1T_2T_3T_4$	D_1	D_2	D_3	$T_1T_2T_3T_4$	D_1	D_2	D_3	$T_1T_2T_3T_4$	D_1	D_2	D_3
1 1 1 1	.6000	.1000	.3000	2 1 1 1	.1538	.1538	.6923	3 1 1 1	.8852	.0656	.0492
1 1 1 2	.7849	.1279	.0872	2 1 1 2	.3358	.3284	.3358	3 1 1 2	.9219	.0668	.0114
1 1 1 3	.1923	.0385	.7692	2 1 1 3	.0262	.0314	.9424	3 1 1 3	.6522	.0580	.2899
1 1 2 1	.8547	.1453	0.0000	2 1 2 1	.4949	.5051	0.0000	3 1 2 1	.9297	.0703	0.0000
1 1 2 2	.8574	.1426	0.0000	2 1 2 2	.5006	.4994	0.0000	3 1 2 2	.9312	.0688	0.0000
1 1 2 3	.8305	.1695	0.0000	2 1 2 3	.4495	.5505	0.0000	3 1 2 3	.9168	.0832	0.0000
1 1 3 1	.0385	0.0000	.9615	2 1 3 1	.0044	0.0000	.9956	3 1 3 1	.2647	0.0000	.7353
1 1 3 2	.1525	0.0000	.8475	2 1 3 2	.0196	0.0000	.9804	3 1 3 2	.6183	0.0000	.3817
1 1 3 3	.0050	0.0000	.9950	2 1 3 3	.0006	0.0000	.9994	3 1 3 3	.0431	0.0000	.9569
1 2 1 1	1.0000	0.0000	0.0000	2 2 1 1	1.0000	0.0000	0.0000	3 2 1 1	1.0000	0.0000	0.0000
1 2 1 2	1.0000	0.0000	0.0000	2 2 1 2	1.0000	0.0000	0.0000	3 2 1 2	1.0000	0.0000	0.0000
1 2 1 3	1.0000	0.0000	0.0000	2 2 1 3	1.0000	0.0000	0.0000	3 2 1 2	1,0000	0.0000	0.0000
1 2 2 1	1.0000	0.0000	0.0000	2 2 2 1	1.0000	0.0000	0.0000	3 2 2 1	1.0000	0.0000	0.0000
1 2 2 2	1.0000	0.0000	0.0000	2 2 2 2	1.0000	0.0000	0.0000	3 2 2 2	1.0000	0.0000	0.0000
1 2 2 3	1.0000	0.0000	0.0000	2 2 2 3	1.0000	0.0000	0.0000	3 2 2 3	1.0000	0.0000	0.0000
1 2 3 1	1.0000	0.0000	0.0000	2 2 3 1	1.0000	0.0000	0.0000	3 2 3 1	1.0000	0.0000	0.0000
1 2 3 2	1.0000	0.0000	0.0000	2 2 3 2	1.0000	0.0000	0.0000	3 2 3 2	1.0000	0.0000	0.0000
1 2 3 3	1.0000	0.0000	0.0000	2 2 3 3	1.0000	0.0000	0.0000	3 2 3 3	1.0000	0.0000	0.0000
1 3 1 1	.0698	.2326	.6977	2 3 1 1	.0090	.1802	.8108	3 3 1 1	.2784	.4124	.3093
1 3 1 2	.1543	.5029	.3429	2 3 1 2	.0247	.4822	.4932	3 3 1 2	.3710	.5374	.0916
1 3 1 2	.0118	.0471	.9412	2 3 1 3	.0013	.0322	.9664	3 3 1 3	.0857	.1524	.7619
1 3 2 1	.2272	.7728	0.0000	2 3 2 1	.0467	.9533	0.0000	3 3 2 1 *	.3981	.6019	0.0000
1 3 2 2	.2312	.7688	0.0000	2 3 2 2	.0477	.9523	0.0000	3 3 2 2	.4035	.5965	0.0000
1 3 2 3	.1968	.8032	0.0000	2 3 2 3	.0392	.9608	0.0000	3 3 2 3	.3554	.6446	0.0000
1 3 3 1	.0020	0.0000	.9980	2 3 3 1	.0002	0.0000	.9998	3 3 3 1	.0177	0.0000	.9823
1 3 3 2	.0089	0.0000	.9911	2 3 3 2	.0010	0.0000	.9990	3 3 3 2	.0749	0.0000	.9251
1 3 3 3	.0002	0.0000	.9998	2 3 3 3	.0000	/0.0000	1.0000	3 3 3 3	.0022	0.0000	.9978

Table 6

1 = test not used
2 = answer "yes" to the test
3 = answer "no" to the test

Computer time GE-265 -- 26 sec.

T_1	T_2	T_3	T_4	Conditional Probabilities of D_1	D_2	D_3
+	+	+	+	1.	0.	0.
+	+	+	-	1.	0.	0.
+	+	-	+	1.	0.	0.
+	+	-	-	1.	0.	0.
+	-	+	+	0.05	0.95	0.
+	-	+	-	0.04	0.96	0.
+	-	-	+	0.001	0.	0.999
+	-	-	-	0.	0.	1.
-	+	+	+	1.	0.	0.
-	+	+	-	1.	0.	0.
-	+	-	+	1.	0.	0.
-	+	-	-	1.	0.	0.
-	-	+	+	0.4	0.6	0.
-	-	+	-	0.35	0.65	0.
-	-	-	+	0.07	0.	0.93
-	-	-	-	0.002	0.	0.998

Table 8

In Table 7 we see that the vector (T_1^-, T_2^-, T_3^+) (the line with a *) is not sufficient to know whether the patient has D_1 or D_2. But, as it is easy to check (in Table 8), the super-sets $(T_1^-, T_2^-, T_3^+, T_4^-)$ and $(T_1^-, T_2^-, T_3^+, T_4^+)$ are not themselves sufficient. In this case all it is possible to say is that the patient doesn't have the disease D_3.

Moreover we can check $\{T_1, T_2, T_3\}$ is a basis p-filtering set. The time and the cost for this set are respectively 10 and 14.

If necessary, all that has been done under the definitions given above could be modified without any difficulties when defining p-exact filtering sets as follows:

we say that a set S of tests is a p-exact filtering set if it is a p-filtering set and if, for any realisation S^o of S, there exists a disease D such that

$$P(D/S^o) \geq 1 - \varepsilon ,$$

$$P(D'/S^o) \leq \quad , \forall D' \neq D.$$

Obviously $\{T_1, T_2, T_3\}$ is not a p-exact filtering set, nor $\{T_1, T_2, T_3, T_4\}$. In this case we can say that the diagnosis table is incomplete; complete if the set of all tests is a p-exact filtering set.

As done in Mattei-Faure-Yacoub (1) we may use the entropy

$$H(D\ ;\ S) = \sum_{j=1}^{3} P(D_j/S) \log \frac{1}{P(D_j/S)}$$

and say that a subset S is a p-filtering subset if it is a p-filtering subset as defined above, and if the information quantity given by any super-set of S is negligible for example, with $S = \{T_1, T_2, T_3\}$, this means that

$$|H(D\ ;\ T_1, T_2, T_3) - H(D\ ;\ T_1, T_2, T_3, T_4)| \leq \varepsilon'$$

(ε' chosen sufficiently small).

Comments

An important case is when a test T is not used to determine a specific disease D.

In the logical model, we may consider that the disease D has two distinct forms, $D^{(1)}$ and $D^{(2)}$ corresponding to T = 1 and T = 0. The method can be modified to this case without any difficulty.

In the probabilistic model it is sufficient to define P(T/D) = 0.5.

Applications

The logical formulation has been used to solve several diagnosis problems (see Begon-Tremolieres-Sultan (1), and Begon-Tremolieres-Sultan-Gouault (1)). Possible applications other than in the medical area include:

-- test of electrical networks

-- chemical tests

-- credit scoring

-- evaluation of advertising campaigns.

REFERENCES

1. F. Bégon, R. Trémolières (1). Elaboration d'un modèle logique du diagnostic, Bulletin de l'IRIA, No. 9, October 1971.

2. F. Bégon, R. Trémolières, C. Sultan (1). Presentation et illustration d'un modele du diagnostic à propos d'un exemple hématologique, Journees d'informatique médicale, IRIA, 78 - Rocquencourt, France, 1972.

3. F. Bégon, R. Trémolières, M. Gouault, C. Sultan (1). Pratique du Diagnostic Logique, Hambourg Congress, 1972.

4. M. Mattei, J. Faure, M. Yacoub (1). Aide de l'informatique au diagnostic médical. Application en toxicologie clinique d'urgence, Rev. Inf. Med. Vol. 2, No. 2, June 1971.

Bibliography

1. N.E. Betaque, G.A. Gorry (1). Automating Judgemental Decision Making for a Serious Medical Problem, Management Science, Vol. 17, No. 8, April 1971.

2. A. Bouckaert (1). Reflexions sur le diagnostic médical. Les campagnes de dépistage, Rev. Informatique Med., 2(3), 243-251, 1971.

3. J. Donio, J. Minery, M. Tenenhaus, J. Ulmo (1). Le problème Bayesian. Application des modèles bayesiens sequentials au problème de l'aide a l'action médicale, SEMA, Rapport de Recherche No. 28, Janvier 1970.

4. A.S. Ginsberg, F.L. Offensend (1). An Application of Decision Theory to a Medical Diagnosis-Treatment Problem, IEEE Trans. on Systems Science and Cybernetics, Vol. SSC-4, No. 3, September 1968.

5. L.B. Lusted, (1). Introduction to Medical Decision Making, Charles C. Thomas, Publisher, Springfield, Illinois, U.S.A. (1968).

6. A.A. Popov, V.A. Shul'ga, V.M. Yanenko, A.D. Krislov (1). Automation of the medical diagnostic process, IFIP Congress 71, August 1971.

7. D.H. Stimson (1). Utility Measurement in Public Health Decision Making, Management Science, Vol. 16, No. 2, October 1969.

8. R. Trémolières (1). Les méthodes d'exploration pour les programmes en nombres entiers, European Institute for Advanced Studies in Management, W.P. 72-10, March 1972.

9. R. Trémolières (2). Une théorie logique de la décision en horizon fini, Congress on Praxeologie, Prague, April 1972.

D. WARNER NORTH
FRED L. OFFENSEND
CHARLES N. SMART

PLANNING WILDFIRE PROTECTION FOR THE SANTA MONICA MOUNTAINS: AN ECONOMIC ANALYSIS OF ALTERNATIVES

This paper demonstrates how formal quantitative methods can be used in planning fire protection policy for a particular geographical area. As an example, it presents an economic analysis of three wildland fire protection policies for the Santa Monica Mountains, northwest of Los Angeles. The alternatives are augmented programs of (1) fire prevention and initial attack; (2) fuel break systems; and (3) fire-resistance roofs and brush clearance around homes. Quantitative models are developed to determine the various fire-related costs and losses that would be incurred under each policy. It is shown that the most attractive alternative from the standpoint of minimising total expected cost plus loss to society is the implementation of fire resistant roofs and brush clearance around homes. Economic incentives are suggested for carrying out such a policy.

10. Planning Wildfire Protection for the Santa Monica Mountains: An Economic Analysis of Alternatives
by D. Warner North, Fred L. Offensend, Charles N. Smart

Introduction

What is the best way to protect a large area against fire over a period of many years? Possible answers to this question quickly translate into specific issues that must be settled by local government and fire agency officials: What size budget should be given to the fire agency? What types of building codes, ordinances, and zoning restrictions should be imposed on the citizens? Decisions of this order are generally made by officials on the basis of experience, intuition, and precedent.

Fire protection agencies must compete for scarce tax dollars and are increasingly faced with demands to justify expenses and demonstrate cost-effectiveness of agency programs. Social unrest and environmental concerns sometimes cloud the complex technical issues involved in modern fire protection. Special pressure groups employ articulate spokesmen and lobbyists to protect their interests, while citizens are becoming progressively more skeptical of the expertise of local officials. As a result, citizens, fire officials, and local governments all tend to be dissatisfied with the way in which fire protection policy is established.

The need for comprehensive strategic planning has never been more evident, but formal quantitative planning has not yet become a common method for generating fire protection policy. In order to choose a policy that meets the needs of a particular geographical area, a common basis must be established for use in evaluating alternative policies. An obvious approach is to use economics. The best policy for fire protection should be the one that minimises the overall total of the cost of carrying out the program plus the losses due to fire. The idea of using economics as a basis for planning policy is not a new one; it was proposed as early as sixty years ago.[1,2,3] The main obstacle to putting this economic principle of minimising cost plus loss into practice has been the difficulty of measuring the costs and losses. However, this difficulty can be overcome, as we shall illustrate in this paper. We shall describe an economic analysis of alternative wildland fire protection policies for the Santa Monica Mountains.* The analysis has provided a useful framework for fire protection planning in that area. We believe that the same approach could be equally useful in many other aspects of fire protection policy planning.

Wildfire in the Santa Monica Mountains

The Santa Monica Mountains northwest of Los Angeles face one of the most serious wildfire threats of any area in the world. Weather, vegetation, and the large population in the area all combine to make the occurrence of large wildfires an inevitable phenomenon. Hot, dry winds called Santa Anas are common during several months of the year. Dessicated by several days of Santa Ana winds, the native brush and grass present a nearly unbroken carpet of tinder-dry fuel. A thoughtless recreation-seeker, an arsonist, or a curious child with matches can cause the spark that ignites a disaster. Driven by Santa Anas, which sometimes gust to speeds of 80 miles per hour, a fire can quickly reach a size and intensity that is virtually impossible to control. Despite the best efforts of firemen and modern fire-fighting technology, severe fires burning

---

*. The analysis and conclusions presented in this paper are based on a study performed by the authors for the U.S. Forest Service. For statistical purposes, the study concentrated on the Los Angeles County and City portion of the Santa Monica Mountains. This area covers 150,000 acres and does not include any land in Ventura Country. The results of the study are documented in SRI Report No. MSU-2275 entitled, "Decisions Analysis of Fire Protection Strategy for the Santa Monica Mountains: An Initial Assessment", June 1973.[4]

under Santa Ana conditions usually burn until the wind dies or the fuel is exhausted. Often the fire stops only when it reaches the Pacific Ocean.

If the Santa Monica Mountains were an isolated wilderness area, the problem of wildfire might be of relatively minor concern. However, this is not the case. With its ocean and mountain views, the area provides a highly desirable residential setting for Los Angeles City and County. The area has approximately 100,000 residents and 30,000 homes, many of which are worth more than $100,000. Fires that escape initial attack often burn thousands of acres and destroy hundreds of homes. The Wright fire of 1970 swept a distance of 8 miles to the Pacific Ocean in a matter of 6 hours. The Bel Air fire of 1961 destroyed nearly five hundred homes in one of the most elegant suburban areas of Los Angeles.

Future fires in the Santa Monica Mountains threaten the destruction of millions of dollars worth of property. This threat is a serious concern for residents, property owners, fire protection agencies, and the insurance industry. A number of proposals have been made to improve fire protection in the Santa Monica Mountains, ranging from greater emphasis on fire prevention activities to the fireproofing of structures. We shall describe below how a comprehensive evaluation of these proposals can be carried out.

An Outline of the Approch

To develop and economic framework for planning fire protection policies for the Santa Monica Mountains, we begin by identifying promising alternatives to the present system of fire protection. We then assess the costs and losses for the present system and for each of the alternatives.

We use average annual values to characterise the various elements of cost plus loss. Historical records show that fires in the Santa Monica Mountains tend to occur intermittently, with several years sometimes elapsing between major fires. Operating expenses for fire suppression activities, however, normally accrue on an annual basis. Therefore, all elements of cost plus loss are assessed on an average annual basis. In addition, all cost and loss elements are converted into monetary terms so that all fire protection alternatives can be evaluated on a common scale.

Evaluating Alternatives for Improved Fire Protection

There are three basic approaches to reducing the losses caused by fire:

1. Limit the number of fires that occur.

2. Given that a fire occurs, limit its extent.

3. Given a fire of specified extent, limit the damage it causes.

Any comprehensive approach to planning fire protection must include all three approaches, but proposed changes in policy can often be placed in one of these three categories. In the Santa Monica Mountains, the following policies were proposed as possible alternatives to the present protection system:

1. Limit the number of fires reaching significant size by establishing better programs for prevention and initial attack.

2. Limit the extent of large fires by implementing an extensive fuel break system as an aid to suppression efforts.

3. Reduce damage by making homes and structures more fire-resistant. This might be achieved by installation of fire-resistant roofs and/or clearing nearby brush.

To evaluate properly whether any of these alternatives would be preferable to the present system, it is necessary first to calculate the expected annual cost plus loss for the present system. Certain elements of cost plus loss, such as the present loss due to housing destruction, are based primarily on the historical statistics on fire destruction in the area. Other elements require more judgemental inputs from experts familiar with the Santa Monica fire problem. By summing the elements, an overall figure for the cost plus loss of the present system can be obtained. This result then provides the starting point for the evaluation and comparison of alternative fire protection policies.

The Present System of Fire Protection

Fire records show that during the period from 1953 to 1970 there were 21 fires of over 100 acres each in the Santa Monica Mountains.* (Records are incomplete for the years prior to 1953.) Collectively, these fires burned an average of almost 6,000 acres per year. They destroyed a total of 831 homes, or 46 homes per year on the average. During this same period, the number of homes in the area averaged 23,000. Extrapolating this rate to the 30,000 homes presently in the area, the average burn rate is presently estimated at 60 houses per year.

Elements of Cost and Loss

To provide a basis for comparing the various protection alternatives, we must determine the total cost plus loss of the present system of fire protection. The total cost plus loss is subdivided into the following six categories: housing-related losses; watershed damage; loss of life; disruption of public services and damages to aesthetics, wildlife, and recreation; brush fire-fighting costs; and program implementation costs.

a. *Housing Related Losses*

Housing values in the Santa Monica Mountains range from $20,000 to more than $200,000 per home. According to real estate agents and tax officials, the average house is worth $65,000. Of this amount, $25,000 is for land. Subtracting the value of the land and adding $10,000 for the value of insured contents gives and average insured value of $50,000 per house for the structure and contents. The average burn rate of 60 houses per year therefore results in an average insured housing loss of $3 million per year.

Uninsured Losses: In addition to insured losses, homeowners face a number of other losses that are not covered by insurance. Some of these are tangible and others are intangible, (such as the psychological trauma of fire), but they must all be considered in any comprehensive evaluation of fire protection policy. As an approximation, we assume that the average homeowner faces a potential uninsured loss of $10,000. That is, he would forego a potential payment of $10,000 to avoid the uninsured consequences of fire. For the 60 homes destroyed annually, this value results in an average uninsured loss of $600,000 per year.

Insurance Sytems Costs: The cost of providing fire insurance must also be included in the assessment of structural losses. We have already accounted for the portion of the insurance cost that covers insured losses. But, in addition to payable claims, insurance premiums are set to cover industry overhead and profit. This factor, the insurance system cost, must be included in our assessment of total cost plus loss because it is a cost of protecting the area from fire. Fir insurance premiums in California are currently set so that the insurance systems costs average 81.8 percent of

---

* Our analysis focuses on brush fires of over 100 acres since fires smaller than this size have generally had a negligible effect on the average rates of burned acreage and structural destruction from wildfire.

expected claims. Therefore, in addition to the $3 million in insured losses, the assessment of housing-related losses must include $3 million x .818 = $2.45 million for insurance systems costs.

Allowance for Partial Destruction Losses and Loss of Other Improvements: Insurance experts estimate that losses for outbuildings and other kinds of improvements, as well as losses to dwellings damaged but not totally destroyed, average one-third of the losses for completely destroyed dwellings. We shall therefore increase the above losses by one-third, or $2.02 million, to give a total figure of $8.07 million per year for average overall housing-related losses.

b. *Watershed Damage*

Fire-related watershed damage can result in mud slides, downstream flooding, debris damage, soil erosion, and sedimentation of reservoirs. Unlike other areas of Southern California, however, Santa Monica Mountains generally have not suffered serious watershed damage. The subsurface soil in the area is relatively stable and the creeks empty directly into the ocean, thereby eliminating much of the debris damage.

To calculate the expected level of watershed damage, we assume that this damage is directly proportional to the number of acres burned. Based on discussions with watershed experts in the area, we used a value of $100 per burned acre for watershed damages. The average burn rate of 6,000 acres per year therefore results in expected watershed damage of $600,000 pe year.

c. *Loss of Life*

Very few people have been killed by wildfire in the Santa Monica Mountains. As an upper bound on the loss of life, we assume that one person is killed every two years in the Santa Monica Mountains. Using a value of $300,000 per statistical human life* gives an expected loss of $150,000 per year.

d. *Disruption of Public Services and Damages to Aesthetics, Wildlife, and Recreation:*

Disruption of public services, such as communications and transportation, and damages to aesthetics, wildlife, and recreational facilities are other measurable losses attributable to fire. These losses have not always been included as direct costs of fire, but they are real losses to the affected segments of society and must be included in any comprehensive analysis of fire protection policy. As a first approximation, we assume that each homeowner in the Santa Monica Mountains would be willing to pay $15 to $20 per year to avoid each of the two categories of loss. This gives a total cost of approximately $500,000 per year for disruption of public services and $500,000 for damage to aesthetics, wildlife, and recreation, or a total of $1 million damage for both categories.

e. *Brush Fire-Fighting Costs*

With the assistance of Los Angeles City and Country Fire Deparment officials, a prelinary assessment was made of the routine costs of maintaining a brush fire-fighting force in the Santa Monica Mountains. Our estimate of these costs is $1 million per year. In addition, the marginal suppression costs of fighting unusually difficult fires are estimated to average $200,000 per year. These marginal costs represent the opportunity losses that accrue because certain activities could not be carried out while men and equipment were committed to the fire.

---

* This figure is an upper bound on the value used by several governmental agencies to approximate the societal worth of a statistical human life.[5]

f. *Program Implementation Costs*

Another economic factor to be considered is the cost of implementing an alternative fire protection program. The cost associated with the present system has already been represented by the brush fire-fighting costs above.

Total Cost Plus Loss Under the Present System of Fire Protection

Table 1 summarises the various cost and loss elements for the present system of fire protection. The table shows that the present system has a total annual cost plus loss of $11.02 million per year. Although many portions of this total do not appear in fire department records (for example, the insurance systems cost), the total does represent the expected annual cost plus loss for the present system of fire protection in the Santa Monica Mountains. This amount will serve as a benchmark in evaluating other alternatives for fire protection.

Alternative 1: Limit the Number of Large Fires

Since the large fires occur when an ignition is followed by failure to control the fire at a small size, the incidence of large fires can be reduced by activities that reduce ignitions and/or increase the effectiveness of initial attack.

TABLE 1

EXPECTED ANNUAL COST PLUS LOSS FOR THE PRESENT SYSTEM OF FIRE PREVENTION

(Thousand of Dollars)

Housing Related Losses			
Insured Housing Losses	$3,000		
Uninsured Losses	600		
Insurance Systems Cost	2,400		
Total Loss from Destroyed Houses		$6,000	
Loss of Other Improvements		2,020	
Total Housing-Related Losses			$8,070
Watershed Damage			600
Loss of Human Life			150
Disruption of Public Services and Damages to Aesthetics, Wildlife and Recreation			1,000
Brush Fire-Fighting Costs			
Maintenance of Brush Fire-Fighting Capability			1,000
Marginal Suppression Costs			200
Program Implementation Costs			0
Total Cost Plus Loss			$11,020

Rather than begin by modeling the effectiveness of a new fire prevention program such as this, we will first compare the expected reductions in loss achieved under a hypothetical program to the cost of implementing that program. Then, if the reductions appear great relative to the cost, we will undertake the necessary modeling efforts. For example, if a new prevention or initial attack program were to reduce the average rate of housing losses and burned acreage by 10 percent, it would reduce present losses and marginal suppression costs by 10 percent, or $1 million per year. (There would be no significant change in the cost of maintaining the regular brush fire-fighting capability because there would still be brush fires to fight.) The resultant $1 million reduction in cost plus loss is approximately equal to the annual cost of adding fifty firemen to the area.

If the addition of fifty firemen to the area or the implementation of other measures costing the same amount of money could generate at least a $1 million reduction in annual fire losses, addition emphasis on limiting large fire incidence would be justified. Many fire experts, however, feel that activities of prevention, detection, and initial attack are already being vigorously pursued and that the total cost plus loss for fire protection would not be significantly reduced by further efforts to reduce large fire incidence.

Alternative 2: Establish an Area-Wide Fuel Break System

A second approach to wildland fire protection is to reduce the extent or size of the large fires that do occur. Fuel breaks are useful in assisting such suppression efforts. Fuel breaks are large strips of land on which fuel volume has been reduced to permit fire suppression crews better access, increased effectiveness, and greater safety. Experts have estimated that a fuel break one-half mile wide would stop the head of wildfire in about 50 percent of the cases.

An extensive fuel break system was considered for the Santa Monica Mountains shortly after the devastating 1970 fire season. If the system had been implemented throughout the entire area, if would have covered 30,000 acres, including several strips of land one-half mile wide. Acquisition costs are estimated at $150 million or $15 million per year on an annualised basis (assuming a 10 percent discount rate and an indefinitely long amortization period). Taxes foregone would have been $3.7 million per year. Construction and maintenance costs would have totaled an additional $2 million per year, yielding an overall annual cost of approximately $20.7 million per year for the fuel break system.

If an average 50-percent reduction in fire size and losses is assumed because of the fuel break system, expected fire losses and marginal suppression costs would have been reduced by $5 million per year. The fuel break system, however, is not economically attractive, because the $5 million savings in fire losses and suppression costs are more than outweighed by the $20.7 million cost of the system. On the other hand, if use of the land could be obtained for free, total cost plus loss for fire protection would drop to $8 million per year, because $18.7 million would be saved in acquisition costs and taxes otherwise foregone. Compared with a cost plus loss of $11 million per year for the present system, the fuel break alternative is attractive only if use of the land can be obtained for little or no charge.

Alternative 3: Require Brush Clearance and Fire Resistant Roofs

A third approach to wildland fire protection is to reduce the damage caused by the large fires that do occur. Unlike much of Southern California, the Santa Monica Mountains suffer relatively minor watershed damage. As shown in Table 1, most of the current fire losses in the Santa Monica Mountains are due to the destruction of homes.

There are several steps that can be taken to protect individual homes when large fires burn through an area.[6] For the Santa Monica Mountains the most important steps are the installation of fire resistant roofs and the clearance of dense flammable brush from the area immediately around houses. Other steps include fireproofing eaves and windows, using sprinkler systems, improving access to homes, and training citizens in rudimentary fire-fighting techniques.

The effectiveness of fire-resistant roofing and brush clearance was well documented in the Bel Air fire of 1961. During this fire 484, or 22 percent, of the 2,204 houses in the exposed area were destroyed. Most of the houses destroyed had either poor brush clearance, flammable wood roofs, or both. Of the 105 exposed houses that had wooden roofs and brush less than 10 feet from the house, 57 were destroyed -- a destruction rate of 54.3 percent. Of the 151 exposed houses that had fire-resistant roofs and brush clearance of at least 100 feet, only one was destroyed -- a destruction rate of 0.7 percent. The two destruction rates differed by a factor of approximately 80. The destruction rates for other categories of roof types and brush clearance are given in Table 2.

TABLE 2

DESTRUCTION RATES FOR BEL AIR FIRE

Brush Clearance (feet)	Approved Roofs	Unapproved Roofs
0- 30	67/275 = 0.243	158/319 = 0.495
30- 60	13/239 = 0.054	104/363 = 0.286
60- 100	2/118 = 0.016	28/195 = 0.144
Over 100	1/151 = 0.007	31/210 = 0.148

Source: Los Angeles City Fire Department Records

The Bel Air statistics, compiled by the Los Angeles City Fire Department, are virtually the only substantial data available that correlate destruction rate with roof type and brush clearance.* If we assume that these statistics are representative of the expected destruction patterns for future fires in the Santa Monica Mountains, then we can calculate the effect of implementing the programs of roof conversion and brush clearance.

---

* After our analysis was completed, we had an opportunity to survey the damage of the Rolling Hills fire of 22 June 1973 with representatives from the Los Angeles County Fire Department and the Insurance Services Office. The fire burned 897 acres and destroyed 12 homes south of Los Angeles, mostly in the city of Rolling Hills. The Terrain, vegetation and density of homes in the burned area were similar to many parts of the Santa Monica Mountains.

Table 2-a gives the destruction rates of the Rolling Hills fire by roof type and brush clearance. Although not as many homes were exposed in the Rolling Hills fire as in the Bel Air fire, the destruction rates for the two fires were remarkably similar. The largest category of houses in the Rolling Hills fire, those with wooden roofs and 100 feet of brush clearance, had about the same destruction rate as the equivalent category in the Bel Air fire (13.1 percent versus 14.8 percent). If the Bel Air statistics had been used to predict the expected destruction for the Rolling Hills fire, the prediction would have indicated destruction of 13.3 houses, compared to the actual destruction of 12 houses.

TABLE 2-a

DESTRUCTION RATES FOR THE ROLLING HILLS FIRE

Brush Clearance (feet)	Non-Wooden Roofs	Wooden Roofs
0- 30	*	1/1 = 1.000
30- 60	0/7 = 0.000	0/9 = 0.000
60-100	0/3 = 0.000	2/6 = 0.333
Over 100	1/29 = 0.034	8/61 = 0.131

*No houses in this category. Source: Data compiled by the authors.

As an example of this type of calculation, we will consider a program that requires brush clearance for 100 feet around all structures. Table 3, obtained from insurance industry records, gives the present distribution of houses in the Santa Monica Mountains by roof type and brush clearance.

TABLE 3

PRESENT DISTRIBUTION OF HOUSES IN THE SANTA MONICA MOUNTAINS BY ROOF TYPE AND BRUSH CLEARANCE CATEGORY

(fraction of houses in each category)

Brush Clearance (feet)	Approved Roofs	Unapproved Roofs	Total
0- 30	0.035	0.004	0.039
30- 60	0.072	0.014	0.086
60-100	0.090	0.026	0.116
More than 100	0.536	0.223	0.759
Total	0.733	0.267	1.000

Source: Brush Surcharge Books maintained by Insurance Services Office.

The table shows that approximately 76 percent of the homeowners have 100 feet of brush clearance. If all the homeowners were to clear their brush to at least 100 feet and leave their roof type the same, then the distribution of houses by roof type and brush clearance would be as in Table 4.

TABLE 4

DISTRIBUTION OF HOUSES IN THE SANTA MONICA MOUNTAINS IF ALL HOUSES WERE TO HAVE 100-FOOT BRUSH CLEARANCE

(Fraction of houses in each category)

Brush Clearance (feet)	Approved Roofs	Unapproved Roofs
0- 30	0.000	0.000
30- 60	0.000	0.000
60-100	0.000	0.000
More than 100	0.733	0.267

The expected destruction rate for a fire is calculated by multiplying the destruction rate for each roof-type/brush-clearance category by the fraction of houses in that category and then summing over all categories. Using the data in Tables 2 and 4, the expected destruction rate, given fire exposure and universal implementation of 100-foot brush clearance is:

$$E(\text{destruction rate}|b_{100}, f) = (.007)\ (.733) + (.148)\ (.267)$$

$$= .045$$

where

$E(x|y)$ = expected or average value of the quantity x, given y

b_{100} = universal implementation of 100-foot brush clearance

f = exposure of area to wildfire

Thus, if a fire were to sweep through the Santa Monica Mountains after the brush was cleared to a distance of at least 100 feet from all structures, the expected percentage of houses destroyed would be 4.5 percent. This compares with a 22 percent rate for the Bel Air fire itself. The difference is due to the improvements in brush clearance and a lower proportion of wooden roofs than existed in the Bel Air area in 1961.

To determine the average annual burn rate, we must multiply the 4.5 percent rate, which is conditional on a fire going through the area, times the probability that the area will be exposed to wildfire in a given year. Historical statistics show that an average of 6,000 acres are burned every year in the Santa Monica Mountains. Since there are 150,000 acres in the area, this means that there is an average cycle time of about 25 years between major fires in a particular area. Housing tracts, however, generally receive more fire protection than undeveloped areas, and so we shall assume an average cycle time of 30 years between major fires in a particular housing area. This gives a probability of exposure in any one year of 1/30 or .033. Thus, the expected annual burn rate for the case of 100-foot brush clearance is:

$$E(\text{annual burn rate}) = .045 \times .033$$

$$= .0015$$

Since there are 30,000 homes in the area, the average number of homes that would be lost to wildfire each year under a program of 100-foot brush clearance is:

$$E(\text{homes destroyed}) = .0015 \times 30{,}000$$

$$= 45 \text{ houses}$$

In other words, implementation of a program of universal brush clearance would mean that an average of 15 homes per year could be saved based on the present average destruction rate of 60 houses per year.

The annual cost plus loss with a policy of universal brush clearance is calculated in the same manner as for the present system of fire protection. In this case, however, the damages must be adjusted to reflect the new burn rate of 45 houses per year. The cost of clearing the necessary brush (at an average cost of \$100 per acre per year) must also be included in the calculation. These calculations show that the expected annual cost plus loss for a program of 100-foot brush clearance is \$9.6 million per year, or approximately \$1.5 million less than the present system of fire protection.

Similar computations can be done to determine the effect of either universal implementation of fire resistant roofs alone or the joint implementation of fire resistant roofs and 100-foot brush clearance. For purposes of brevity, the calculations are not given here. They are carried out exactly as for the case just described, except that the distribution of houses by roof type and brush clearance must be revised to reflect the program being considered. The calculations assume that the annualised cost of converting a roof from an unapproved to an approved rating averages \$390 per converted roof (which is the approximate annualised cost of installing a fire-resistant pressure-treated shake roof).

Table 5 compares the present system of fire protection with programs for additional brush clearance and fire resistant roofs. From an economic standpoint, any one of these three fire protection programs would be more attractive than the present system. The most attractive program, however, is the one requiring universal implementation of both fire resistant roofs and brush clearance. This program would reduce expected housing destruction by almost a factor of 10, from the present rate of 60 houses per year to 7 houses per year. Even though it would cost $3.7 million per year to implement the program, the total expected annual cost plus loss would be reduced to $7.6 million per year -- compared to $11.0 million per year under the present system of fire protection.

Our analysis shows that of the three basic approaches to wildland fire protection, the most attractive one for the Santa Monica Mountains is that of requiring fire resistant roofs and 100 feet of brush clearance throughout the area. The large cost of acquiring land for fuel break systems and the comparatively small benefits from additional programs of fire prevention and initial attack make these programs unattractive compared to a combined program of roof conversion and brush clearance.

TABLE 5

COMPARISON OF PROTECTION POLICIES INVOLVING
BRUSH CLEARANCE AND ROOF CONVERSION

Protection Policy	Average Annual Number of Homes Destroyed	Annual Program Cost of Protection Policy (Millions of Dollars)	Average Annual Cost Plus Loss to Society, Including Program Cost (Millions of Dollars)
Present situation: existing roof types and brush clearance	60	$0	$11.0
Existing roof types and native brush removed 100 feet from all homes	45	0.6	9.6
Existing brush clearance and conversion of all wood roofs from unapproved to approved type	21	3.1	8.9
Both brush clearance to 100 feet and conversion to approved roofs	7	3.7	7.6

Sensitivity Analysis

In carrying out our analysis, we made a number of assumptions, including several preliminary value assignments. Sensitivity analysis is useful in determining how sensitive our conclusions are to changes in these assumptions. By changing the values of the different variables and recalculating the cost-plus-loss sums we can determine the range of conditions over which our conclusions are valid. Since this paper does not permit a detailed analysis of all variables, we will examine one variable in detail and summarise the findings of the other sensitivity analyses.

The fire cycle time is defined as the average time between major fires in a particular area. This is an important variable because it gives the probability that a house would be exposed to wildfire in a given year. Based on historical statistics, we assumed a cycle time of 30 years, which gave a probability of exposure of 1/30, or 0.033. We wish to examine how important this assumption is because, for example, if major fires are actually very rare events, then the present system of fire protection may be fully adequate.

Figure 1 summarises the cost-plus-loss calculations for cycle times ranging from 10 to 70 years. The graph was constructed by calculating the probability of exposure for a particular cycle time and then determining what the average annual losses would be for each of the more attractive fire protection plans. All other variables in the analysis were kept at their initial values. The graph shows that as long as the cycle time is less than about 50 years, the most attractive alternative is the one requiring fire-resistant roofs and 100 feet of brush clearance. If the cycle time is less than our assumed value of 30 years, then this alternative is even more attractive than is indicated in our analysis. On the other hand, if the average time between major fire exposures is greater than 50 years, the most attractive alternative is that of requiring brush clearance alone.

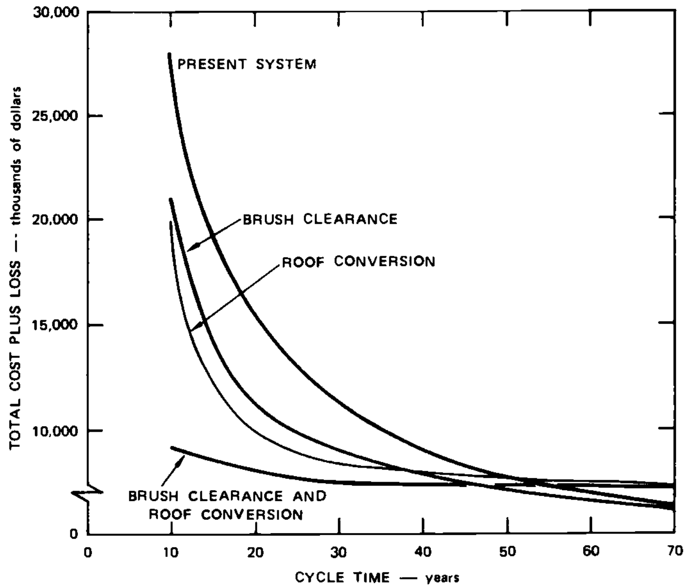


Figure 1 SENSITIVITY TO FIRE CYCLE TIME

Sensitivity studies for all of the variables are documented in our expanded report.[4] The studies confirm that the most attractive alternative for a wide range of conditions is that of universal implementation of fire-resistant roofs and 100 feet of brush clearance. For example, even if the actual probabilities of destruction were only 70 percent of the statistics observed in the Bel Air fire, the most attractive alternative would be the same. The value used for uninsured losses has no effect on changing the preferred alternative. Only if the average value of a house plus contents were less than $28,000 per house (compared to our nominal value of $50,000 per house) would the preferred alternative change to that of requiring brush clearance alone.

Incentives for Implementing Programs of Brush Clearance and Roof Conversion

Ordinances that require 100 feet of brush clearance and fire-resistant roofs for new houses are in effect in Los Angeles City and County but there is a high degree of noncompliance, as shown in Table 3. Only 54 percent of the existing homes presently have both approved roofs and adequate brush clearance. Some of the reported noncompliance is a result of the difference between what the insurance industry considers to be an "approved" roof and what the ordinances require to meet the standard of "fire-resistant". Pressure-treated shake roofs have passed certain Underwriter Laboratory tests and are considered fire-resistant by both the insurance industry and city and country ordinances. Dip-treated shake roofs are considered fire-resistant by the city and county ordinances, but not by insurance underwriters.

An alternative to strict dependence on legal ordinances is to provide homeowners with an economic incentive for carrying out roof conversion and brush clearance. In fact, one type of incentive system already exists: the brush surcharge that is added to the fire insurance premium for homeowners in the Santa Monica Mountains and certain other areas in Southern California to reflect the threat of wildfire. The surcharge typically varies from $80 to $800 per home per year, depending on roof type, brush clearance, and fire protection class. However, the differences in the surcharge rate are not proportional to the destruction rates observed in the Bel Air fire. For example, the surcharge for a $50,000 home with 100 feet of brush clearance is $100 if the house has an unapproved roof and $80 if the house has an approved roof (for a house in an area with fire protection class 4B). As shown in Table 2, however, houses in the latter brush-clearance/roof-type category are only 5 percent as likely to be destroyed if a wildfire occurs.

Several purposes would be accomplished by revising the surcharge schedule to reflect more precisely the likelihood of destruction. It would serve as an economic incentive to homeowners to clear their brush, install fire-resistant roofs, and take other measures to protect their property from wildfire. The revised schedule would also eliminate present financial inequities: no group of policy holders would be subsidising other groups. Finally, the premium schedule would clearly indicate to the homeowner the actual difference in risk he faces. If a homeowner realises that his house with an untreated shake roof is twenty times more likely to burn in a brush fire than his neighbour's house that has a fire-resistant roof, he might readily decide to replace his existing roof with a fire-resistant type.

Conclusions

Our analysis, supported by sensitivity studies, has shown that from a societal standpoint the most cost-effective means for protecting the Santa Monica Mountains from wildfire is for all houses to have fire-resistant roofs and 100 feet of brush clearance. If this plan were implemented, structural losses could be reduced by almost a factor of ten, decreasing the present average burn rate of 60 houses per year to a rate of only 7 houses per year. Although wildfire is inevitable in the Santa Monica Mountains, homeowners, developers, local government officials, insurance executives, and bankers should be made aware that certain protective measures can greatly reduce wildfire losses.

Ordinances requiring fire resistant roofs and 100 feet of brush clearance presently exist in the area. The level of complicance, however, is low. Economic incentives such as an equitable brush surcharge could provide the necessary motivation for improved compliance. The present surcharge rates appear to be unrealistically low for houses with poor brush clearance and wooden roofs. The brush surcharge should be set to accurately reflect the actual likelihood of destruction for the different categories of brush clearance and roof type.

Our analysis depended heavily on the statistics from the Bel Air fire, because these were the only data available for our purposes. We do not wish to imply that the destruction rates for that one fire should be adopted uncritically as the basis for evaluating wildland fire protection policy. Rather, we believe that fire protection policies should be based on a careful assessment of how local destruction rates depend on roof type, brush clearance, and other relevant factors.

In evaluating policies of housing protection, we focused on programs of brush clearance and roof conversion. However these methods are not the only ways that homes can be protected from wildfire. If future research efforts find that there are less costly or more aesthetically attractive ways of achieving the same protection, then they should be seriously evaluated. Research to determine ways of lowering the cost of fire-retardant shingles should be considered. Other worthwhile investigations might include evaluating the role of ornamental shrubbery in the spread and confinement of wildfire and the feasibility of pruning native brush as an alternative to clearing it.

As our study has shown, the cost-plus-loss approach provides a simple and logical framework for evaluating wildland fire protection policies. Because it is easy to use and provides information that can be translated into practical policy decisions, we feel this methodology can be of great benefit in many areas of fire protection planning.

REFERENCES

1. Dubois, C. "Systematic Fire Protection in the California National Forest," U.S. Forest Service pamphlet (U.S. Government Printing Office, Washington, D.C., 1914).

2. Arnold, R.K. "Economic and Social Determinants of an Adequate Level of Fire Control," doctoral dissertation, University of Michigan, Ann Arbor, Michigan (1950).

3. Davis, K.P. Forest Fire: Control and Use, (McGraw-Hill Book Company, New York, New York 1959).

4. Howard, R.A., North, D.W., Offensend, F.L. and Smart, C.N. "Decision Analysis of Fire Protection Strategy for the Santa Monica Mountains: An Initial Assessment," Preliminary report, U.S. Forest Service PSW Grant No. 6, SRI Project MSU-2275, Stanford Research Institute, Menlo Park, California (June 1973).

5. Sagan, L.A. "Human Costs of Nuclear Power," Science, Vol. 177, pp 487-493, (11 August 1972).

6. Christian, W.J. "The Effect of Structural Characteristics in Dwelling Fire Fatalities," Fire Journal, pp. 22-28 (January 1974).

JOHN STRINGER

OPERATIONAL RESEARCH AND PUBLIC POLICY

The final chapter in this collection provides a global review of the progress of analytical methods in public policy formulation in the United Kingdom. Particular emphasis is given to the organisational and behavioural aspects of implementation and some directions for future research are suggested.

11. Operational Research and Public Policy
by John Stringer

*"Operational research has been slow in penetrating the broad social and economic sphere, where policy decisions affect the life of the ordinary man profoundly."**

That was in 1962, since when many more people have been engaged in Operational Research (OR) in government and various organs of the public sector. Whether this is evidence of faster penetration into the field of public policy is more questionable. It is the aim of this paper to examine critically and hopefully constructively whether OR can have such an impact, and how.

Machinery for Policy-Making

Many innovations have been made in recent years in the machinery by which policies are formed in the United Kingdom on matters of public concern. Among the changes have been:

- merging of Ministries into big Departments of State, with the intention of improving policy coordination;

- Public Expenditure Survey (HMSO 1961) (Mackenzie 1969) (Heclo and Wildavsky 1974), by which Parliament is able to consider the future costs implicit in current policy intentions;

- the Department of Economic Affairs (short-lived but, for example, its regional planning structures have survived);

- policy planning units in Government Department; as recommended by the Royal Commission on the Civil Service (HMSO 1968);

- Programme Analysis and Review (Heclo and Wildavsky 1974) (the titles and contents of these longer range reviews of policy options are not made public, but the selection of subjects appears to be a process of bargaining between departments);

- Green Papers, introduced in the 1960's as a vehicle for setting out options for public discussion;

- special political advisers to Ministers, providing them with a channel of advice separate from that of the permanent officials;

- reorganisation of local government, health services, the water industry, etc, into larger units deemed capable of sustaining adequate planning and policy analysis expertise;

- hiving off from government of agencies whose function was considered to be to execute, rather than to create policies, eg: the Manpower Services Commission;

- the Central Policy Review Staff, introduced to provide the Cabinet with a channel of advice and comment across the whole field of policy;

---

* Tavistock Institute and Operational Research Society, 1962.

- corporate planning concepts introduced into local government at the time of reorganisation, with the intention of countering the tendency to functional separation;

- structure plans and regional strategic studies (both vehicles for the expression of policy choices at the interface between central and local government);

- planning agreements between government and industry.

Joining the European Community, Devolution, and other constitutional changes have been argued in terms of their effect on the quality of policy-making. Outside government many academic bodies and institutes are declaring themselves to be 'in the policy business' (Sharpe, 1975) and new ones, such as the Centre for Studies of Social Policy have come into being. 'Policy studies', 'policy analysis' and 'policy sciences' (to cognoscenti, not all the same) are terms to conjure with nowadays.

Similar things are happening the other Western countries. In Canada a new Institute for Research on Public Policy is funded from central and provincial government and private sources. This followed a report by Ritchie (1971) which was an important statement of the need for, and character of, public policy research. In West Germany a Commission on Economic and Social Change has, since 1971, been engaged in studies of the policy implications of trends and aspirations in that society. It has tripartite sponsorship of employer and employee organisations and the social science community.

In the Netherlands a statutory Scientific Council for Government policy has been established and given power to call for meetings with the Cabinet to discuss action on its advice. An early step was to 'call the government's attention to the need for more co-ordinated preparation of long-term policies for European affairs, the structure of the economy, energy, the labour market, education, public health and social welfare and the distribution of income, wealth, power and knowledge'.

These examples, and others, show that whilst approaches vary there is widespread belief that policy-making can be improved. Presumably it is felt that better policies will emerge, and thus a better future than it would otherwise be.

Such propositions cannot be subjected to straightforward empirical test since, in an uncertain environment, good policies would not lead to good results every time, nor bad policies to bad results. Only in rare and rather atypical cases can controlled experiments be done to compare alternative poicies. It is still less feasible to make a comparative evaluation of policy processes. So, as Dror (1973) puts it 'the proof of the pudding is in the cooking' (or the recipe). With so many cooks rebuilding the policy kitchen, one can but wonder what is happening to the broth.

Why All These Changes?

What lies behind this concern with policy and attempts to change the way policy is formed? Might it be no more significant than a temporary phase of governmental fashion? Does it result from the availability of more sophisticated techniques? Could it be part of a long-term trend perhaps highlighted by the desire of incoming governments to make visible changes? Will it disappear under the attempts now being made in many countries to halt or reverse the growth of the governmental sector?

The view taken in this paper that concern with policy-making is based on a special present-day need which arises from genuinely felt uncertainty about where society is going, and that this need will not be satisfied easily or quickly. It follows from this view that anything which purports to improve policy-making and to reduce uncertainty ought to be critically examined.

Uncertainty is nothing new. At a personal level we deal with it by holding on to values and beliefs and by setting the uncertainty against a background of the familiar and reliable. If uncertainty comes a bit at a time we have ways of coping with it. We may, for instance, off-load it on to other broader shoulders.

The effect of a concatenation of many uncertainties is of a different quality. It can lead to disorientation and disturbed behaviour. When it hits society the sense of confusion and the felt need for some stable elements become very strong. To the extent that widely held and accepted norms and values are not available to provide this stability, and to the extent that authoritarian imposition of rules and guidelines is unacceptable, it becomes necessary for deliberately negotiated policies to try to deal with more and more aspects of life.

Anxiety is, I believe, one of the main contributory causes of the current concern with public policy and for the continuous enlargement of its domain. As examples of these anxieties:

- realisation that the finite nature of natural resources may be a real constraint on the growth, or even the continuance, of industrialised society. Industrial and employment policies which seem to contradict one another and a general lack of knowledge as to what the alternatives might be, add to the uncertainty;

- the implications of the desire for greater participation and for individual development and self-actualisation, contrasted with increasing reliance on government through social services and other interventions;

- the loss of familiar guidelines implicit in the removal of social norms which could previously be taken for granted. These include adjustment to sex equality, changed attitudes to marriage, attitudes to law and order, acceptance of authority. Many of these uncertainties result from reforms which in themselves are liberal and well-meaning.

If this list actually succeeds in expressing the nature of significant anxieties it is fortuitous since the subject has not been thoroughly investigated. (It would be a useful contribution to policy research if there were a continuously updated source of such information).

If there were a consensus on societal goals and on the values which underlie them, of if those who did not share the consensus lacked influence (or did not care to use it), then the need for policy would be minimal. The coordinating factor would be the system of shared values. This in turn would be sustained by the fact that individual decisions, wherever and whenever, would be guided by similar principles, and be sufficiently correlated in their effect for no substantial interference to arise. Each agency would take its decisions on the reasonble assumption that the others were not too far from the consensus and that their ideals were not also undergoing major change. Thus there is minimal need for explicit policy where:

- there is consensus on goals and values;

or - expectations are few and simple;

or - abundant resources enable simultaneous pursuit of different goals;

or - decisions on one matter cannot greatly affect the options open to other decision-makers;

or - there are not seen to be any significant alternatives amongst which to choose.

The normal historical condition has been for one or other of these to apply to large parts of the public domain - or the appearance of stability has been sufficiently credible for there to be no great demand for wide-ranging policy analysis to precede critical policy choices. But no longer.

The public policy system has, however, grown up on the princople that issues could be tackled on their merits, as they arose.

We are moving into an age of increasing turbulence, in the word of Emery of Trist (1965), or of loss of the stable state as Schon (1971) puts it. I find it easier to work with a concept of multiple uncertainties and to regard policy-making as the attempt to manage or adapt to them.

A further stage, hopefully not to be reached but conceivable, is where there is insufficient social coherence for even minimal policy consensus to be generated, no reason to expect that any policy will survive long enough to have any effect, and where fundamental processes of conflict management have broken down.

Thus the present time lies between one extreme where policy-making (and hence policy study and analysis) is relatively simple but is not urgently needed, and the other extreme where it is not even possible. The motivation for greater attention to policy formulation would be salient at the present time whether or not there were adequate methodologies of policy analysis to serve it. If my diagnosis is correct, the anxieties which lie behind the contemporary emphasis on public policy will not be allayed by reorganisations and administrative reforms and such panaceas are only likely to create disillusionment. Nor is it it self-evident that the anxieties will be allayed by the use by government of more facts, better analyses and more sophisticated models.

The need to question the usefulness of OR as a way of meeting the demand for policy analysis should now be evident. If the demand is for someting to act as a surrogate for a missing value consensus, then the use of any techniques which implicitly assume that certain values exist and are acceptable will not do. The conditions under which OR can make a valid contribution to policy-making must be spelled out more carefully.

Policy and Policy-Making

At this point I must try to clarify the term 'policy'. This unfortunate necessity arises since the normative study of policy (although not the study of particular policies) is relatively recent and the concepts have not settled down.

It is easy to believe, but deceptive, that the nearer one gets to working at the top levels of an organisation, the nearer one is to working on policy. This is implicit in the position adopted by Dror (1973) who

> "regards policy analysis are 'preferization' of policy options according to the utilities of 'legitimate decision makers'".

This stance begs some questions, and may be contrasted with that of Lindblom (1965, 1968) who regards policy-making as a process of partisan mutual adjustment, i.e. as one in which interest groups with different utilities vie with one another to move things to their own advantage. This view is essentially an incrementalist one, although in his later writings Lindblom recognises that planned change of a substantial rather than a marginal nature is conceivable and the possibility of it should be taken into account. Whether, and in what way, the partisans engaging in the mutual adjustment process are representative of the range of interests in society and whether any are systematically excluded are important questions of course, but they cannot be pursued here (see Playford 1968).

Heclo and Wildavsky (1974), in a study of public expenditure decisions, give support to the view of policy-making as a matter of political bargaining adding that a lot of it takes place in a small, and relatively closed, system. They describe the strong cultural influences within the small group of government officials who engage in the sophisticated game of allocating public expenditure in Britain. Career motivation and how officials acquire esteem within the system are, they suggest, major influences.

Willson (1969) described the policy-makers of Britain as a set of about 350 politicians and civil servants. He sees these people as engaged in the hurly-burly of top level decision but having no time to think. The implication is that a major part of the making of policy, including the generation of ideas and the pressures bringing a subject to the surface at any particular time, are taking place elsewhere. The distinction between policy-making in a total sense, and policy decision in the limited sense is a significant one.

Dahl (1958) criticises the popular interpretation which sees policy influence as located in ruling elites. By defining operational criteria by which a hypothetical elite could be demonstrated to be such, he challenges whether what are often believed to be ruling elites are in fact so. I conclude that one should not jump to conclusions about how policy is made and by whom.

A more specific concept of policy is used by Friend et al (1974) in work which deals largely with relations between central and local government. They refer to a policy as a statement, often emanating from the centre, intended to give guidance on a defined class of more tactical decisions. A useful concept deriving from this model is one of 'policy stress' i.e. a situation where several separate guidelines, all applicable to a single instance, give conflicting indications.

The similarities and differences between 'a policy' as a deliberated guideline, and a 'value' as a guideline resulting from a more general cultural process are interesting. Policy is intertwined with values: sometimes tending to create or to adapt them; sometimes substituting for them; sometimes ignoring them; sometimes placing one above another; and sometimes responding to them by articulating their application to particular types of case. Policy and values are as hen and egg.

The different usages of the term policy can be confusing. Partly this arises from the differences between a descriptive and a prescriptive viewpoint. From the descriptive point of view I prefer to think of policy-making as a continuous social process in which certain aspects of society become singled out for attention, and certain stances become adopted in relation to them. This avoids the tautologies that can occur when definitions assume the prior existence of a special class of people who are 'the policy-makers'.

Within this neutral descriptive frame of reference one can then locate for prescriptive purposes; particular areas of policy; the formal and informal systems within which they are formed; and possible interventions in these systems. OR is an example of such and intervention, as would be an organisational change or, say, a spontaneous community action.

Although public policy-making is not co-terminous with the problem of allocating public expenditure they are strongly connected. Apart from the direct connections, certain expenditures may unwittingly create values or expectations of wider significance than the immediate object of the expenditure would suggest. Thus, increased expenditure on social services has probably contributed to a shift in expectations and values; as an example recent instances of cruelty to children have been treated as breakdowns not of private responsibility, but of official responsibility. Such a shift of values, involving the concepts of freedom and responsibility, could obviously spill over into other fields and create further pressures on the public domain. The difficulty of taking such effects into account in the original expenditure decisions will be obvious, and some supposedly rational methods may only exacerbate the problem.

The concept of objectively assessed need has been influential in the rational approach to social services policy. A great deal of research has been directed at measurement of need but it is not always appreciated how relative such measurements can be. For instance, in the case of meals on wheels and other publicly provided meals for the handicapped and the elderly, estimates of need derived from conscientious and competent research studies made at various times between 1958 and 1973 increased exponentially from about 6 million a year to 300 million a year over that period.* Over the same period, provision rose from 1.2 million to about 27 million. It is difficult to avoid the conclusion that 'need' is as much a value judgement as an objective fact.

A recent government circular on reductions in social service expenditure indicated that research should be the first thing to go. Is it surprising when the expectation has been created that research in this field inevitably leads to increased expenditure?

Discussion of policy and policy research solely in terms of public expenditure tends to lead to a formulation in terms of the allocation of marginal monies between mutually exclusive classes of expenditure and this is reinforced by the nature of the official game already referred to. It is another example of the incrementalism which is so prevalent in policy-making processes. The creative redefinition of policy problems and the examination of more substantial** innovations which might offer release from the incremental treadmill should also be provided for and it is not easy to do this within a framework geared too closely to concepts of public expenditure allocation.

The nature of public policy-making can be further examined by drawing a series of contrasts as follows:

A contrast between policy-making and policy decision has already been implied. Those working on policy problems inside an organisation tend to adopt the view that conflicts of interests can be resolved by referring further up the hierarchy. With the more open-system view of policy-making being used in this paper, it is only in special instances that this method of conflict resolution is available. Most top level policy decisions are based on ideas which have come from elsewhere. Seminal ideas are often regarded as socially deviant at the time of their first emergence and those associated with them treated accordingly. In policy-making, why does a particular issue arise? why does it arise in this form? and so on, are important questions; whereas in the top level process of policy decision, answers to these questions can easily be taken for granted.

Public policy can be contrasted with private policy. In the latter 'it is not our policy to ...' is an effective argument stopper, whereas in the realm of public policy it is more likely to start and argument. Expressions like 'in the public interest' which imply that there exists a general social utility, are in fact only used when the intention is to put somebody down and to set aside his utilities as irrelevant. The essential question in matters of public policy is <u>whose</u> costs? and <u>whose</u> benefits? The trend of thinking on participation and the social responsibility of industry are bringing these questions into the private industrial sphere and thus bringing that more into the realm of public policy-making also.

The next contrast is between policy and decision. The following summarises the argument of Bauer and Gergen (1968). Decision making as understood by psychologists, decision theorists, etc., assumes a single decision making unit with a single set of utilities. This unit has a range of options and knows their consequences, intends to make a rational selection, and is able to do the sums. In policy formation these assumptions are violated. However, it is quite usual in discussing policy issues, not to recognise this: thus hindering understanding by diverting attention away from what actually takes place. Using the decision model suggests that policy-makers ought

---

* I am grateful to my colleague Michael Norris for pointing this out.

** The practical significance of the distinction between marginal and substantial innovation for the development of a strategy of public policy research, is well brought out by Chevalier and Burns (1975). I return to this point later in the paper.

to behave in a way which is in fact inappropriate to the situation they are in. Bauer and Gergen's argument obviously has a lot to do with the applicability of OR to policy.

The rational and the political aspects of policy-making are often contrasted, but political behaviour is not ipso facto irrational. For example, the tendency of hospital management to be conducted as a political process has often been criticised but in a situation of such uncertainty, it is natural for the doctors, for instance, to try to maximise their ability to handle whatever the future may bring, i.e. to seek and retain power. If that were the whole story, it would be entirely rational behaviour for them to behave politically.

It is appropriate to conclude this section of the paper by expressing the view, common to many writers, of the increasing importance of interdependence in planning and policy making. This has many causes. Each move which brings another aspect of personal life into the realm of public policy, creates further possibilities of anomaly and inconsistency and further demands for redress or extension. Another factor is the trend for organisations to define their function at higher and higher system levels (Schon 1971). Thus, what once were Ministries labelled as transport, housing, etc., are now the Department of 'the Environment'. A colleague was recently engaged in a study, not of hotels, but of 'the hospitality industry'. In whatever way a public or private redefines the business it is in, the higher level definition is likely to intersect with the definitions adopted by others.

Because of this characteristic of interdependence, a very central set of questions is 'how much co-ordination? what sort? and how?'

Co-ordination

The importance of the amount and means of co-ordination as a factor in policy-making, together with the differences between the latter and decision making, make this a useful focus for exploring the relation between OR and policy. The options in regard to co-ordination each likely to be associated with a particular OR methodology, would include:

- enlarging the span of interests of single organisations, or creating over-arching policy or supervisory bodies;

- identification of policy problems independently of current boundaries;

- large scale models, data bases, or analytic capabilities;

- development of the perceptive and analytical capabilities of individual policy actors;

- development of inter-organisational joint problem solving abilities;

- reducing reliance on formal co-ordination, eg: by allowing price mechanisms to operate.

Organisational Change

It is always easy to criticise the left hand for not knowing what the right hand is doing and since inter-dependence is a growing feature of policy-making, the search for more (and by implication, better) co-ordination is a recurrent theme in administrative reforms. However, such organisational shuffles are necessarily limited to present perceptions of the inter-dependence of various activities and may thereby tend to delay adaptation to future inter-dependence.

Organisational change is traumatic and takes a long time to settle down. It cannot be contemplated frequently. Moreover there is not much scope left for creating ever larger agencies in the hope of better co-ordination. Thus, despite a radical appearance, such changes tend to be geared to incremental changes in policies, and to execution, and to be biased against more far-sighted, and possibly more creative perspectives.

Reorganisations of this kind have, however, often provided the occasion for introducing OR as, for example, on the creation of the Civil Service Department in 1968. Soon afterwards a major OR study was commenced concerned with the dispersal out of London of parts of the Civil Service, on a scale which would have to include policy-makers as well as the executors of policy. This study, in which the Institute of Operational Research has the privilege of participating, has been published (HMSO 1973) but I want to ponder here on some of its implications.

The Dispersal Study

The underlying model was of discrete 'block of work', each fairly homogenous and connected to other blocks by communication links of greater or lesser strength. The effect of dispersing a sub-set of the blocks would be to incur the disbenefits associated with stretching some links. An algorithm enabled 'good' sets of candidates for dispersal to be selected and the cost and 'link damage' consequences to be estimated. (Turner et al 1970).

From the point of view of the present paper, the following points arise:

- it is implied that present patterns of communication are valid over the indefinite period during which the dispersal moves will influence the conduct of affairs;

- no weight was given to the fact that dispersal would make all subsequent reorganisations more difficult to achieve;

- the longer range uncertainties were further set aside by discounting;

- dispersal could affect the career paths of Civil Servants perhaps making moves between different areas of work less likely. Any effects this might have on the experience and viewpoints brought to their work could not be taken into account in the calculation;

- similarly disassociation, over time, from "south eastern" views and attitudes could not be considered. It may be noted that different departments would be dispersed to different parts of the country;

- it was not found possible to examine alternative patterns of government e.g. one having a stronger regional bias, with many functions represented at every region and less national uniformity of treatment of given areas of policy. Such a pattern might either have the advantage of giving about anyway as a result of political pressures for devolution;

- the measures of link strength used in the model did not give any special weight to policy-making as distinct from execution, except to the extent that policy involves more highly-paid officers and frequent meetings;

- although a major purpose was to alter the regional distribution of types of employment, enquiry revealed how little was known of the effects on a region of moving jobs to it. Nor were alternative means of achieving such effects considered.

On the other hand, few organisational changes can have been so well studied at the design stage. It is interesting to contrast with the reorganisation of the National Health Service in England and Wales, which has been based on some untested hypotheses about the nature of consensual management, and about the technical feasibility of comprehensive health planning. Specifically, the dispersal work:

- won a considerable achievement in being accepted as a fair basis for negotiations between interested parties with strong views and skill in presenting them;

- the model did not try to produce an optimum solution but was designed to be used inter-actively, by responsible officers in the course of the negotiations;

- the study recognised and took account of the interests of the people concerned, and examined the many aspects of home life whose interaction with the career aspects could be very significant.

I am forced to observe that the pressures and the circumstances under which this study was undertaken led, probably inevitably, to the use of a model and data in such a way as to reinforce the present pattern of organisation and communications. By using the status quo as data it helped to ensure that substantial changes from the status quo did not enter the set of options and that any such changes that did occur would be due to fortuitous effects outside the scope of the model.

Nevertheless, I do not conclude that the study should not have been done nor that its methodology is useless. Doing the study gave a great deal of insight into the working of the government machine and most of the short-comings arise from the fact that this was a one-off job and was not conceived as part of a continuous, on-going programme aimed at maintaining a body of scientific knowledge to back up administrative changes of various kinds. The dispersal model, if associated with methods continuously for detecting changes of the actual and prospective trend of government work or the emergence of new areas of policy concern, could have become a focal point of important scientific work of practical significance. It could have helped to provide the better understanding necessary for an anticipatory, adaptive, approach to the continuous review of government machinery.

If policy-making and administration are not to be badly caught out by changes in the socio-economic environment, then they must be able to adapt in tune with a changing society. Ideally, working perspectives would be selected from a set as diverse as the range of possible variations in the socio-economic environment itself.

The following is an account of a project which had the aim of steering a governmental machine away from automatic reliance on the current definitions of the boundaries of policy problems.

Problem Structuring

The setting was the design of arrangements for long-range planning in a Western country having a federal constitution. There was little tradition of interchange of personnel and ideas between departments of government, and sectoral policies were therefore developed independently of each other. Scharpf (1971) described the situation as one of 'negative co-ordination' where exchange of information and views occurs after a policy has been worked out by the responsible section of a government department and not before. Operationally, this means that there is no joint perception of the problem boundary, no joint development of options, and no joint discussion of the criteria of choice. On the other hand, it is an inexpensive, and often rapid and decisive way of getting things done.

As the time horizon extends, however, the short-comings of negative co-ordination become more apparent. Moreover, western societies are increasingly reluctant to give automatic acceptance to what 'they' have planned for 'us' so there is increasing risk that if policies are developed from limited perspectives, they will repeatedly have to be returned to the drawing board.

'Positive co-ordination' would mean the coming together of those responsible for two or more subjects, accordingly to a re-structured (and often ad hoc) definition of a policy problem. Together they would develop options, consider criteria and examine constraints and inter-dependences. Positive co-ordination is expensive, bearing in mind the time it takes for people from different backgrounds to discover how to work together creatively. The number of possible pairs or larger multiples that might be brought together in this way is combinatorially large. The question was, therefore, how to select the clusters of interests to be brought together.

Two areas of policy-making ought to be in the same 'cluster' if the policy options in one area were likely to have significant impacts in the other; or if options in the two areas conflict, are mutually incompatible or mutually reinforcing; or if new joint options could be conceived. In other words, the reason for bringing A and B together is that the choices in A are conditional on the options available in (A + B).

The moment a decision is made to cut down the range of options in B the contingency relationship between A and B is altered. It may now be right for them to go their separate ways. It is a matter of judgement whether A and B warrant being brought together or separated and at any time, new factors may arise to alter these judgements. The aim was to make the clustering process as systematic as possible and for it to be dynamic and adaptive.

If this could be achieved, it would serve a purpose which could not be served by shuffling organisational boundaries. The latter could then be designed to meet the continuing purposes of policy implementation and administrative management without being compromised by the different requirements for co-ordination in longer-term policy development.

The model around which a steering mechanism for the policy-development process was designed was as follows. (Fach et al 1972)

Consider a matrix in which rows represent areas of policy and columns areas of impact on the physical, social, economic, etc., environment. Elements of the matrix refer to the impact of policy area i on environmental area j, scaled from 0 to 1. Each element is an answer not to the question 'does i affect j', but to the more subtle question 'to what extent do the alternatives or options still open within policy area i differ (significantly, understood) in their impact upon environmental area j?!

Ideally, in terms of rigorous logic, there would be a balanced set of exclusive policy areas, together giving exhaustive coverage of existing and potential public policies. In practice the definition would be in terms of the responsibilities covered by government departments and public agencies, with allowance for re-definition as particular policy subjects come into focus.

Similarly, an ideal set of environmental areas would be balanced, mutually exclusive, exhaustive, and in sufficiently fine detail. In practice a hierarchically organised set of impacts, which would allow the degree of resolution to be adjusted, was adopted and in the initial experiments we used a level of this hierarchy giving about 100 impact areas.

Impact areas are equivalent to 'social indicators' (Bauer, 1966: Gross, 1967: Terleckyj, 1970). Despite much work, the theoretical base for defining such indicators is shaky (Plessas and Fein, 1972). It has to be recognised that the impact dimension of our model would have to evolve in step with changing perceptions of what is becoming important to society.

Altogether, therefore, the matrix was not conceived as absolute but as a convenient summary of present perceptions and commitments. It tried to make these explicit and to provide a framework for steering the policy-making process, relatively independent of the administrative structure. Setting aside, for the moment, the question of how the entries in the matrix are made, the next step was to derive from the policy-environment matrix, a triangular matrix of similarities between policy areas, a policy-policy matrix. The assumption here is that the superiority of positive over negative co-ordination between any two areas is related to the similarity between their patterns of environmental impact. After some experiments, the elements of the policy-policy matrix were taken to be the normalised co-variances (around zero) between pairs of rows of the policy-impact matrix. This 'similarity coefficient' is 0 when the two policy areas have no common impacts, 1 when their impacts are identical, and is unaffected by environmental areas on which both of them have zero impact (thus avoiding spurious impression of similarity).

Several computing methods are available to pick out clusters of policy areas such that the degree of similarity between members of a cluster is markedly higher than that between members of separate clusters. By adjusting the threshold level at which pairs of policies are deemed similar, the number and size of clusters can be adjusted to match the amount of positive co-ordination activity that can be undertaken within the resources available.

In our experiments the elements of the policy-environment matrix were obtained from panels of officials using the Delphi method (Dalkey, 1969). The intention was to do the same with other expert groups including politicians in central and local government, social scientists, political journalists, etc. The model would be continuously updated as implied by the description given, and re-run from time to time in order to guide the process of forming and re-forming (including disbanding!) joint working parties engaged in developing policies. I have dwelt on this example despite the fact that the work was not completed at the time, because, whether or not the process described was actually used for guiding policy-making, the model brings out some of the problems involved in bringing a 'rational' approach such as OR to bear on policy-making. It emphasises:

- that policy problems, and hence co-ordination problems, are not given, they have to be picked out;
- the importance of selective mobilisation of resources for policy development and analysis;
- the need to change the connections as policies become established, new opportunities arise, and as value shifts occur;
- the need to avoid being trapped by current boundaries of departmental responsibility;
- the essential reliance on judgements;
- the number of such judgements required, and their combinatorial complexity, which suggests that without some explicit steering process, policy-making is likely to be not much different from a random process.*

Practical politics make it hard to change the way in which areas of policy are perceived and tackled, even when the need to change the scope of co-ordination is obvious - see, for instance the careful wording of the Central Policy Review Staff's paper 'A Joint Framework for Social Policy' (HMSO, 1975). There is a lot of inertia (or momentum, if you prefer) in the official policy-making apparatus. Furthermore its mode of survival is, on the whole, to be reactive rather than pro-active

---

* Perhaps that is why the Oxford English Dictionary gives, as one definition *POLICY ... a form of gambling in which bets are made on numbers to be drawn in a lottery.*

in relation to changes occurring in its environment. The following example illustrates how OR sometimes reinforces these conservative characteristics.

BIG MODELS

The models developed in the Department of Energy (Hutber, 1972) to act as a co-ordinating influence on the work of several policy divisions of that Ministry, consist of a predictive econometric model of demand, linked to models of each of the supply sectors - coal, gas, electricity and oil. By 1972, the last of these consisted of equations representing oil as the balancing factor, available in any quantity at a given world price. Thus the model reflected a) current organisation with nationalisation of all the fuel industries except oil; b) world circumstances at the time of its development; and c) the organisation within the Department. It must be expected, therefore, that in application the models tend to re-inforce the perspectives on the energy situation current at the time. Insofar as the model is an instrument of change, such change is likely to be incremental.*

I have heard criticisms on the lines that each of the energy industries has its own models which are technically superior to the corresponding parts of the national energy model. That is barking up the wrong tree. The point is that the model-building effort, and the organisational setting in which it is embedded, does not contain the requisite variety to keep up with the changing socio-economic environment, for example, a) the shift on moral as much as economic grounds, in attitudes about the consumption of natural resources, especially energy; b) social pressures to search for renewable and non-polluting sources of energy; and c) the emergence of oil as a dominant factor in international conflict.

Heaven's above! It's too much to expect any model to be open-ended enough to add factors such as these to the complexities it already copes with. But these ever-changing values, attitudes, and opportunities, are a large part of what energy policy must be about. The UK energy model is important as a means for helping with certain types of investment decision within the nationalised fuel sector. But, in a sense, it brings these decisions, big as they are in money terms, down to the level of routine and thus out of the area of policy-making into that of policy-doing. Whether the model contributes to inertia or whether it helps bring in new options, depends upon policy (or political) processes outside its scope.

This not an attack on the model-builders themselves, and it may be unfair to pick their work to illustrate a general thesis. I merely wish to illustrate that whether and in what way to use models is itself a kind of policy choice. It will be interesting to see what the effect is of the recently announced decision of the Energy Department to make the model available to outside parties interested in the energy sector.

Appraisal of Approaches to Co-Ordination

So far we have been looking at approaches to improved co-ordination, which have been structured and deliberate. Hence, they are likely to be constrained by current organisational boundaries or at least by current perceptions of the need or opportunity for changing them. Were we dealing with the kind of policy-making which falls within the legitimate private domain of an individual or company, there would be no particular cause for comment. Since, however, we are considering the making of public policy, it is of potential concern when the scope and definition of a problem is forced into an inappropriate mould because of organisational constraints, rather than responding to the nature of the underlying public issues.

There is no point in being purist about this. The issues of public policy can be

---

* An analysis of UK energy policy, drawing particular attention to the inertia aspects of the policy system is contained in Chesshire et al, 1976.

seen in a great variety of ways and the existing pattern of official agencies is often a convenient frame of reference for tackling them. Furthermore, if a mismatch persists between the offocial perceptions and those which 'feel real', additional political actors* often emerge with the aim of pressing an alternative point of view. Although 'political' behaviour of this kind may act as a corrective, however, it cannot always be relied upon to act quickly enough, or sensitively enough.

The last statement has implicit within it a stance about the quality of public policy-making. It would be appropriate to bring this out explicitly here by suggesting that the policy-making process is 'better' to the extent that:

- the issues selected for attention are ones which have or will have a profound effect on the lives of ordinary people;

- the processes do not systematically ignore issues of this kind;

- a range of options receives consideration, including the possibility of substantial, as distinct from incremental, change from the status quo;

- significant options are not unwittingly ruled out by the effects of current organisational boundaries;

- conflicts are not created by the process itself, i.e. that the conflicts which arise to be resolved, reflect real differences of values between interest groups.

It would be difficult, perhaps impossible, to operationalise criteria such as these, even if they were universally acceptable. Equally, though, how can one appraise changes, including the use of OR, without some such criteria in mind?

So far we have considered how the co-ordination aspects of policy-making may be improved (or modified, at any rate) by organisational means; including continuous adaptive forms of temporary organisation along with the more conventional 'permanent' forms of re-organisation. The role of OR may, for instance, be as a tool at the organisational design stage; or continually as providing a service to one or more of the actors in the politics of the situation; or as providing a source of information and appreciation neutrally available to several such actors (although rarely to all of them). The feature common to these approaches is that problems and issues tend to be defined in terms so chosen as to match existing organisational boundaries. If they are not so defined, then they are unlikely to attract influential attention.

By the nature of its techniques, by its reliance on quantitative data, and by the historically sanctified aim to be close to an established decision-maker (Waddington, 1973) OR is more likely to be associated with marginal innovation than with substantial innovation. If this is a consciously adopted and understood bias, fair enough. It may not always be realised, however, that opting for the use of OR can also mean adopting the values of a particular form of dynamic conservatism (Schon, 1971).

The single, large, hierarchical organisation and the overarching organisation responsible for policy are two means often invoked for improving co-ordination, but we have seen how the large OR model is naturally associated with such approaches to produce a bias against consideration of substantial change.

---

* The term 'actor' in a political analysis may refer to an individual or some larger group or a whole agency. The usage depends on the issues being examined and the extent to which a group can be defined which has a consistent purposive stance relative to such issues. An OR group may itself become such an actor as protagonist for a certain point of view.

The case against this bias is not that incremental changes, or zero change, are bad in themselves. It is that more substantial changes should at least be seen as on the cards, either because of trends occurring elsewhere or because they offer creative possibilities which are unlikely to be reached as a sequence of incremental decisions.

By long-term building up of knowledge and understanding, OR can make a real contribution to policy, but not by the mere smash-and-grab application of techniques. What needs to be looked at most carefully, however, is how access to this knowledge and understanding is controlled and whether it is used by a sub-set of political actors as a source of power. Overcoming this calls for a process of public scrutiny and challenge, both as to scientific validity and to implicit value assumptions, of OR purporting to illuminate public issues (as distinct from serving sectional interests).

The roles and political behaviour of the various political actors (officials as much as politicians) are often seen as matters of individual personality, but whilst it cannot be denied that such factors do operate, they are associated with structural factors of some importance as the following examples may illustrate.

Individual Co-ordinative Behaviour

In connection with the dispersal of government, a study was made of policy-making behaviour in the Scottish Office in order to see whether there was any marked effect of distance from the centre of governmental activity in London. It was observed that, comparing officials of the same rank in the two places, those in the Scottish Office were likely to have a wider span of current responsibilities than their London counterparts. Moreover, their career paths were likely to have taken them through a variety of functions which, in London, would belong to different Departments which interchange personnel more rarely. As a consequence, it seems that in Scotland it was the more natural for an official to perceive the possibility of connection between the problem he was currently handling, and the questions falling in other areas of responsibility. A more informal mode of communication and a naturally wide-ranging diagnosis of the ramifications of policy issues appeared to result.

It was also observed, however, that the more generalist behaviour in the Scottish Office was dependent to some extent on the existence, in London, of officials with a deeper specialisation who could be consulted if need be - an example of symbiosis between actors in the policy-making process which it would be wrong to ignore.

It is interesting to contrast these observations with 'problem-structuring' as described earlier. In the government concerned in that example, officials made their careers by moving upwards within the narrow functional branch which they first entered. It was partly for this reason that a formal means of scanning for interdependence between areas of policy was seen to be needed.

These observations on individual co-ordinative behaviour were made incidentally during work which had other objectives, rather than in a framework of research on the phenomenon itself. They can only be regarded as indicative, but they are consistent with the findings of another study (Friend, Power, Yewlett, 1974) which concentrated on those behaviours which facilitate inter-agency co-ordination. The setting was the planned expansion of a small town (Droitwich), a project in which the combined effect of the decisions and policies of many official and other agencies could be observed. By plotting the networks of contact maintained by the main local actors in this situation, it was possible to see how at certain phases of difficulty, one or other of the actors would take it upon himself to extend the policy-making system by drawing into the arena people who had not hitherto been involved, but whose involvement now seemed appropriate. This reticulist (i.e. network-forming) behaviour at times utilised considerable skill as well as a range of vision in detecting the need to act in this way. One of the main driving forces was seen to be the state of 'policy stress' which occurred from time to time.

In a sense, what might be regarded as a failure of co-ordination at the central level, whilst being a source of frustration, could also invoke an informal co-ordinative behaviour which could be a source of creativity.

Incidentally the officials concerned had no budget against which to charge this essential activity and had to put it down to such 'respectable' activities as designing road systems. This is a nice reflection of the prevailing values about productive and unproductive work.

In these various examples of the co-ordinative behaviour of individuals, it is possible to observe several factors, including:

- ability to perceive the need to change the boundaries of a policy problem;
- the importance of varied career paths in developing this ability;
- the political skill to bring about appropriate change in the composition of the policy-making system;
- the effect of the setting, and the stresses it gives rise to, as driving forces for the deployment o such skills.

To which might be added:

- the ability (often lacking) of people drawn from different backgrounds, disciplines, professions, etc., to engage creatively in joint problem-solving.

Of course, individual personality is a factor also, but I have tried in this essay to suggest that there are also factors affecting the quality of co-ordination which are consistent and structural and therefore in principle capable of being modelled and hence understood scientifically. I chose the question of co-ordination because evident lack of it is one of the more readily detectable shortcomings of public policy-making, because *prima facie* the need for it expands as complexity and uncertainty grow, and because a variety of approaches is available. These approaches differ in their effects and in the role which OR can play in relation to them. Rather than explore these differences further, however, it is necesary to return and pick up one or two loose ends.

The Unthinkable, and Substantial Change

The conservative and incremental tendencies of policy-making have been referred to several times, as has the observation that OR is often involved in reinforcing these characteristics. It is a pity, in my view, that this should be so. OR was a significant social invention in that in its early days (Waddington, 1973) it provided for the operational commander a scientifically based theory grounded in actual operations. By being able to nandle several variables at once, and by a general spirit of challenging untested assumptions, it could produce ideas which were novel and even radical and support them by evidence and calculation. The ability to offer new options was as significant as the ability to provide a rationale for choice between options that were already known. Unfortunately, a) quantitative analysis and b) being close to a powerful decision-maker, are the characteristics of those early days which have continued to be valued in OR circles. In that the analogy between 'operational commander' and 'public policy-maker' is a false one, it seems to me that these are not the values which should have been carried across as dominant. A more fruitful analogy would have emphasised c) the creative synthesis aspects and d) service to the true analogue of 'commander' - i.e. not a single powerful <u>individual</u>, but a <u>process</u> involving many actors.

Against such a background, the following is a sketch of the kind of thing which might be attempted.

A Cluster of Issues

I have selected the following as salient problem areas which, if looked at together, might generate options that would be unlikely to arise from regarding them as separate problems:

- employment (and unemployment),

- job satisfaction (the quality of working life),

- production of tradable goods and services,

- consumption of natural resources.

Although this is a personal choice, to illustrate a point, clusters of issues of which this might be one, would be brought into prominence by the 'problem-structuring' approach described earlier; or might arise from individual 'reticulist' initiatives.

There are currently intractable aspects to all four problem areas and within each of them policies are currently being pursued which impact on the others. Taking the cluster as a whole could, therefore, suggest options for substantial change (i.e. involving several variables at once) which are more attractive than those that can be conceived of in terms of incremental change (i.e. of one variable at a time).

Evaluation of policy change in any one of these problem areas is made difficult by the connections with the others, so that if (it is a large if!) the cluster could be modelled, it would add a useful element to the relevant policy-making systems. Can we define, then, for the heuristic purpose of generating and evaluating new options, a model consisting of inter-dependent variables extending into all four of the problem areas in the cluster? What factors would the postulated model have to bring into account?

Employment

Policy on employment is more concerned with the avoidance of unemployment than with a clearly formulated positive objective. Unemployment is contemplated with horror and in recent years governments have responded to this by economic policies and by various employment protection measures. They may incidentally have created exaggerated expectations that the government is able to control economic activity to a greater extent than is real and should therefore take responsibility for creating employment wherever unemployed people happen to live and wish to stay. The manner in which such popular expectations arise and are reinforced by the actions of government is of less importance here than the strength of the values held about maintenance and creation of employment. Fear of unemployment persists even though the material hardships are mitigated by social security provisions. It is not, as it once was, a matter of actual starvation. Indeed for some, the material difference between being in a job and being unemployed is small or even negative. In such cases the social stigma of being unemployed is the uppermost consideration.

Employment has become the central social ritual of industrial man (Pym, 1975). In arriving at this ritual position 'employment' has drawn upon, and distorted, the positive values which once belonged to 'work'. A moment's thought will show that work and employment are not the same, for most people they overlap, but neither completely contains the other.

Insofar as employment is the central ritual, to be excluded from it is to be denied access to valued systems of social support, to feel worthless, punished and deterred; a latter-day excommunication. Putting it another way, an important sub-set of the satisfactions and dissatisfactions a person obtains from life, is associated with employment as distinct from work. To incorporate them in a model as though they were synonymous would be to invite misleading conclusions. Phrases like 'the journey to work' show that the distinction is not always made when it should be.

The anthropological concept of ritual is relevant to the question of how unemployment is to be managed. Consider a few of the options.

The position of the school-leaver without a job to go to has been a matter of recent concern because of the waste of human resources capable of useful work and because of the deviance problems which can arise when young people are deprived of the socialising effects of employment. A policy of job creation has been adopted with a view to providing school-leavers with experience of paid jobs on such projects as clear up bits of the physical environment. The reactions, which have been mixed, are interesting: they seem to say that:

- for young people acclimatised to a scruffy urban environment the projects lack point, it isn't a 'real job' to do now what has been left undone for decades;

- not being on the payroll of a 'proper employer' doesn't carry the ritual overtones of employment even though it _is_ work (physically) and _is_ paid for.

Some, at least, of the actors in the employment policy-making system have the objective of maintaining present types of jobs in existing factories faced with closure. From the standpoint of normal economic theory, this represents an inefficient allocation of resources. It has, however, led to the interesting development of the 'workers co-operative'.

The skills of the leading participants in such schemes, deriving from their experience as shop stewards, are in organising production on the shop floor. As production units, and as sources of satisfaction to those involved, some of these experiments seem successful. It is more doubtful whether they can cope with the problems of product development and marketing for instance, but with parallel innovations in these functions, there may be more robust possibilities.

The point here is that, although the success of such experiments may be qualified in their present context, they may nevertheless contain some of the components which in combination with other ideas could create new options, once freed from preconceived assumptions about the organisation of manufacture.

One of the main options of employment policy is the creation of new job opportunities through investment. Little need be said of this here, except to note the resource implications, and especially the energy implications, of the levels of capital intensity normally involved. Schumacher (1973) is the main protagonist of the alternative option of 'intermediate technology'. Although aimed at developing societies this may convey some messages for industrial societies also.

Other principal levers of employment policy operate by constraint, e.g. by making it difficult and expensive for an employer to make people redundant. Whatever else such policies do, they cannot avoid reinforcing the belief that it is employment which is important and work that is incidental.

Other concepts about employment are subject to the process of reinforcement by usage. For instance, that on the whole people have one job (at a time), for one employer, during normal working hours, at a wage determined by an annual round of bargaining. Departures from these norms of stereotypical job are punished (i.e. made more difficult) by, for example, the national insurance regulations, the general rule that 'unsocial hours' justify higher pay, and in numberless other ways. But suppose these assumptions were relaxed and changes made such that people found it better to have two jobs? Would employment carry the same stigma then? Would not mid-career adaptation through re-training be more feasible? Retirement less of a sudden shock?

Rituals have their origins in the necessities of existence. Characteristics of the employment ritual which now seem dysfunctional have been displaced from the circumstanced in which they originally acquired valued status, i.e. the fact that the goods and services needed for even a modest level of survivial required that most people did physical work most of their time. This need is no longer absolute in a world of 'post-scarcity' but it takes time for a corresponding change in the social mores of work, the consequences of which could be far-reaching (Higgin, 1973). Meanwhile the position can be summarised as:

- employment is valued,
- work and employment are not the same,
- the satisfactions (and dissatisfactions) derived from employment relate to its ritual aspects as well as to the work done,
- the values attaching to the ritual aspects are likely to adapt and change evenutally but not without warning (if the signs can be interpreted),
- satisfaction often comes from work outside the employment context (voluntary service, family activity, do-it-yourself),
- it is conceivable that patterns of employment and of the organisation of work could be devised which increase nett satisfaction,
- this is not a zero-sum game. We might all gain.

Job Satisfaction

In recent years increasing attention has been given to the idea that work in the context of employment can be made more satisfying (Wilson, 1973). A common feature of approaches such as job enrichment and job enlargement, is that they seek to meet psychological needs through the design of the work itself rather than by compensatory 'hygiene' measures. Hertzberg (1968) concluded that the content and structure of jobs, rather than the surrounding conditions, were the main factors in motivation and satisfaction. Removing the things which cause dissatisfaction does not itself produce satisfacticn which seems to be related to a different set of factors.

The social and psychological desiderata being pursued in the redesign of work situations have been set down (Thorsrud, 1968). At the level of the individual they involve optimum variety of tasks; a meaningful pattern of tasks that gives to each job a flavour of a single overall task; satisfactory length of work cycle; some scope for the individual to set standards of quantity and quality of production and a suitable feedback of knowledge of results; the inclusion in the job of some of the auxiliary and preparatory tasks; requirement for a degree of care, skill, knowledge or effort that is worthy of respect in the community; the job should make some perceivable contribution to the utility of the product for the consumer.

At the level of the group the objects of redesigning are to provide for 'interlocking' tasks, job rotation or physical proximity where there is a necessary interdependence of jobs; or where the individual jobs entail a relatively high degree of stress; or where the individual jobs do not make a obvious perceivable contribution to the utility of the end product. Where a number of jobs are linked together by interlocking tasks or job rotation, they should as a group contribute visibly to the utility of the product, and the group should have some scope for setting standards and receiving knowledge of results, be able to handle its own internal organisation and have some control over 'boundary' tasks.

Most social scientists working in this field stress the participation of people in the redesign of their own jobs. The reasoning seems to range from the desirability of tapping their knowledge and experience, to a more idealogical commitment to industrial democracy. This kind of participation at the level of the design of the day-to-day working activity is somewhat different from the concept of participation by workers' representatives in policy-making. The latter is becoming more a matter of collectivist power bargaining and hence more concerned with the satisfactions obtained from employment than from those obtained from work.

The aspect of job satisfaction which is often the presenting symptom first calling attention to its significance, is the phenomenon of withdrawal, usually manifested as absenteeism, sickness, or labour turnover.

It is, I think, significant that the many studies of job satisfaction and the schemes introduced have been in the context of incremental change within existing employment situations. A great deal has been learned about the psychological and social satisfactions which people derive from work and employment and this knowledge might usefully be incorporated in a more wideranging model. A diagram may help to summarise the argument so far, Figure 1.

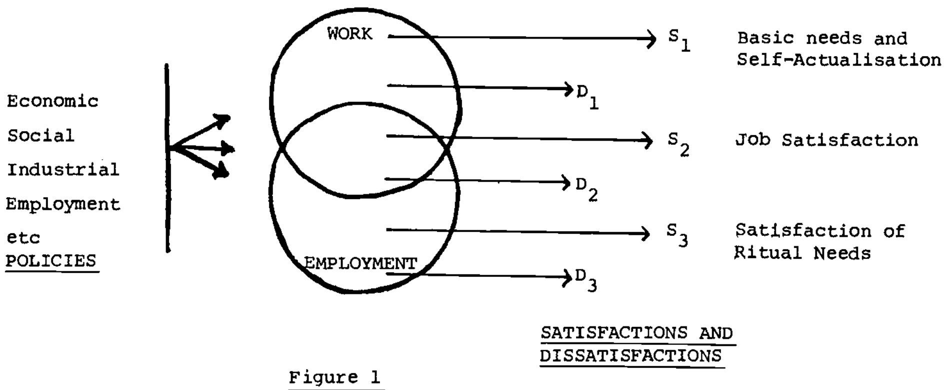


Figure 1

Tradeable Goods and Services

As it stands the diagram misses out what an economist would regard as the whole point of industrial activity, namely the production of goods and services. There are interesting value overtones to be considered here since, to many, production of material products is 'good' and output of 'unproductive' services at best dubious. Bacon & Eltis (1975) draw a distinction of greater economic relevance between tradeable and non-tradeable goods and services. The fact remains, however, that feelings about the value society places on the output of work is a factor in job satisfaction and may be irrational. The service of mending things could, in a context of limited resources, be seen as better than the productive activity of making new ones. As these values adapt there will be a change in the satisfactions derived from work and employment from this cause alone.

As the basic physical necessities of life such as food and shelter are satisfied and we seek what would usually be regarded as higher things, the different value implications of the outputs of various kinds of work become increasingly relevant to public policy and increasingly difficult to comprehend in terms of economics. If we now add to our model the kinds of satisfaction of human needs which derive from the production and consumption of goods and services, it becomes Figure 2. One thing this diagram does not depict is an answer to the question 'whose satisfactions?' Although this situation is not zero-sum, it must sometimes happen that one's gain is another's loss and I have no wish to gloss over the element of conflict which the model must adequately reflect both in its construction and in its use.

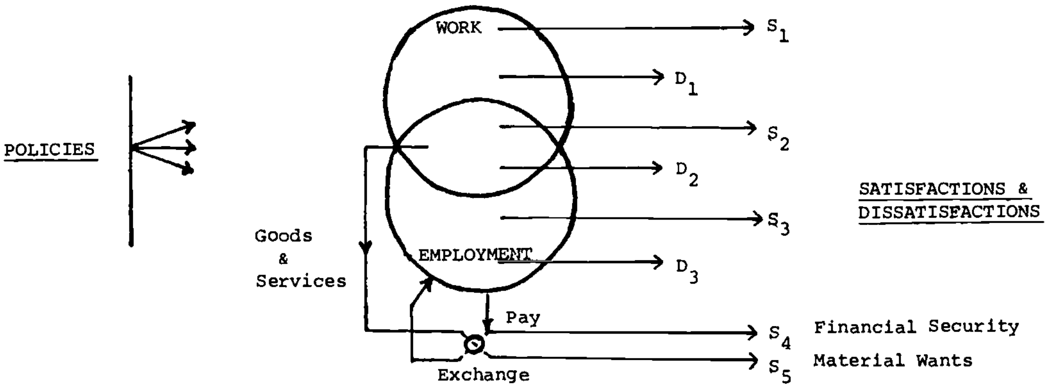


Figure 2

Resources

There are other conflicts. There has been widespread concern that the natural resources to support present levels of industrial activity will run out sooner or later; and fear that it might be sooner. As a crude summary of the vast amount of research and speculation on this point the debate falls between two schools. The 'doom' school fear that man will run out of renewable resources, with disastrous effects. Their views tend to be reinforced by the 'conservatism' and 'back to nature' schools. Economic growth will have to be reversed, they say, with consequences for every kind of policy. The 'cornucopians', by contrast believe that man's ingenuity will find substitutes in good time for anything which becomes scarce. With enlightened economic policies, they say, growth can continue, since economic forces and technological 'fixes' will carry us through. A more modest position, between these extremes, is that provided there is enough energy, and the institutions of society can adapt to making investment decisions which are sufficiently long-sighted, and world shifts in economic power do not lead to major wars, then continued economic growth is feasible and the world system need not collapse.

Energy problems look very different as seen by economists, by politicians, by technologists, according to concepts familiar to their own disciplines and to the different time horizons to which they are accustomed in their thinking. As a matter of public policy energy is a large central affair but this hardly reflects the fact that actions and decisions throughout society have substantial energy consequences. Thus energy is a relevant unit of account by which to incorporate resource questions into our model.

Energy must have a central place in this discussion because it is the ultimate resource. That is, all transformations of materials require energy, and with sufficient of it acceptable substitutes can be found for anything which becomes scarce. Goeller and Weinberg (1975) conclude that mineral resources are adequate provided man finds an inexhaustible non-polluting source of energy; the main problem is how to make the transition from the present state of relative plenty of oil, coal and other resource materials to what they call 'The Age of Substituability' using renewable resources only. The problems they see are social and political; new social institutions would be needed to overcome the fact that the market place optimises short-term advantages, thus inhibiting the transition. Appropriate policy changes are, therefore, contingent upon institutional changes.

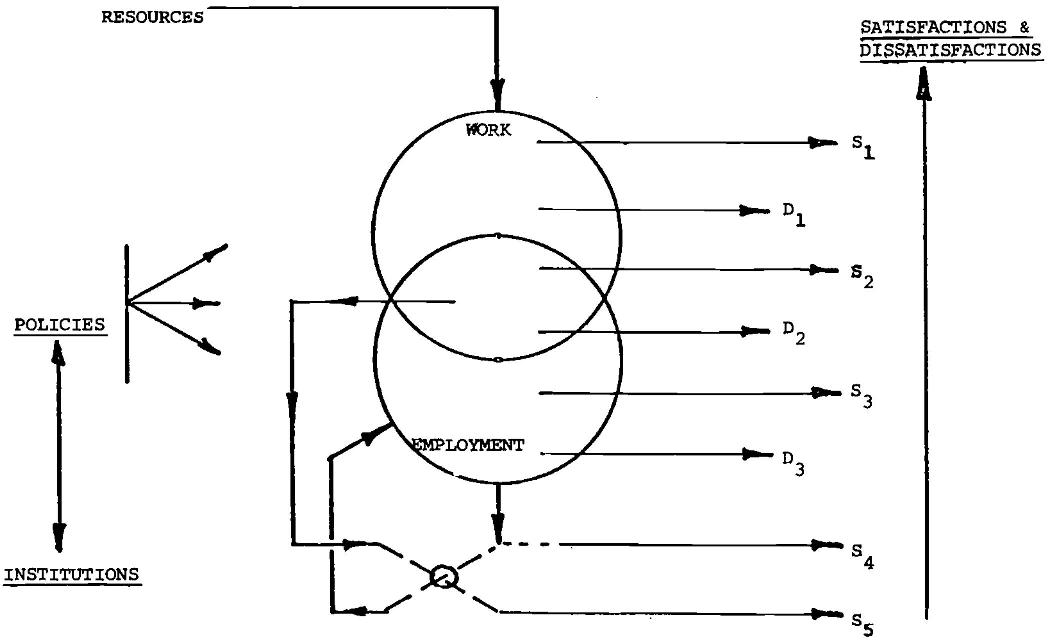


Figure 3

We have now redefined the cluster of problems in terms of a balance between an input of resources and an output in terms of the satisfaction of human needs and desires. This balance is influenced by policy interactions and institutional change. For the purpose of reaching broad conclusions energy will serve as a single numeraire of input especially since a good deal of research on energy accounting (e.g. Wright (1975) and Slesser 1975)) could be used. We are not yet in a position to simplify the satisfaction side of the model, and might never be. Let us assume for the moment, however, that we can.

There is implicit in discussion of economic growth the idea that greater human satisfaction will be, and can only be, obtained via greater production of material goods and hence by the greater consumption of resources, but it may not be so. Maslow (1954) claims that there is a partially ordered hierarchy ranging from the basic physiological needs such a food; shelter and other needs for safety; then needs for love and esteem. The upper part of the range is the drive for individual self-actualisation which he expresses as 'what a man can be he must be'. Maslow has suggested that the attempt to obtain satisfactions further up this hierarchy is made as soon as, but not until, the more basic requirements have been satisfied. In that extravagant consumption of material goods is sometimes a surrogate for the satisfactions to be derived from individual self-actualisation, it could be that the optimum relationship between the degree of satisfaction achieved and the consumption of material resources, is as in the following diagram. Hopefully it will be found to be so.

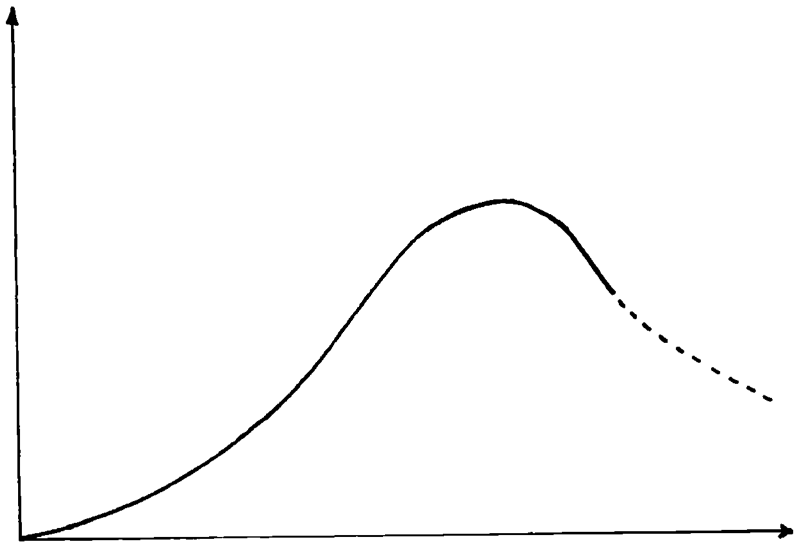


Figure 4

The Model, or, and Policy-Making

I have been referring to a 'model' although what has been described so far would not serve the usual OR purpose of enabling 'what if?' calculations to be made. It should be thought of as a field of enquiry rather than as an all-embracing computable model. The suggestion is that it would be possible to generate and appraise new patterns of axtivity to see whether greater human satisfaction can indeed be achieved at the consumption of less natural resources than is at present the case. I offer it as a conceptual model and as an example of what OR needs to be doing if it is to make a contribution to public policy which is neither:

- restricted to incremental change,

nor - tied to the interests of a single political actor.

The process of change is, of course, usually an incremental one involving individual political actors. OR done in such settings can still be good and useful and make an important contribution to policy-making. But it would be so much better if it were possible to take into account substantial changes either as options for decision or as some of the scenarios of the future against which present options should be appraised.

Conclusion

I do not know what institutional provisions need to be made to correct the incrementalist and other biases of OR, and to re-deploy some of its technical capabilities on the lines suggested. Perhaps that is a subject for another paper.

I can only summarise this one by reiterating that:

- policy-making is an important and complex social process;
- it is misleading to think of policy-making as simply a special kind of decision-making;
- OR might contribute to policy-making in a variety of ways;
- the various stances of OR differ in the extent to which they are linked to policy issues (defined independently), to policy processes, or to particular policy actors, or to the interactions between these perspectives;
- in practice OR tends to be biased towards reinforcing present perspectives, present options, and present influences, i.e. to be dynamically conservative;
- recognising that some of this bias is unconscious, there is a need for complementary work concerned with the generation and appraisal of options for substantial change.

This would mean discovering, for the more difficult field of policy problems, a spirit similar to that which OR had in its early days when it offered a new view of operational problems.

References

1. Bacon, R.W. and Eltis, W.A., 1975. Declining Britain. Articles in the Sunday Times, 2nd November to 14th December, 1975.

2. Bauer, R.A., 1966 (ed) Social Indicators. MIT Press, Cambridge, Mass.

3. Bauer, R.A. and Gergen, K.J., 1968. The Study of Policy Formation. Collier-Macmillan, London. The Free Press, New York.

4. Chessire, J.H., Friend, J.K., Pollard, J.deB., Stringer, J. and Surrey, A.J. 1970. Energy Policy in Britain: a case study of adaptation and change in policy systems. Prepared for a conference at Breau-sans-Nappe. To be published.

5. Chevalier, M. and Burns, T. 1975. A Recommended Strategy for Action-oriented Public Policy Research. Prepared for the Institute for Research on Public Policy, Canada.

6. Dahl, R.A. 1958. A Critique of the Ruling Elite Model. American Political Science Review, Vol. 52, No. 2, pp 463-9.

7. Dalkey, N.C. 1969. The Delphi Method: an experimental study of group opinion. Rand Corporation, Santa Monica, Calif. Memorandum RK-5888-PR.

8. Dror, Y. 1973. Public Policy Re-examined. Intertext and Leonard Hill Books, Aylesbury and New York.

9. Emery, F.E. and Trist, E.L. 1965. The Causal Texture of Organisational Environments. Human Relations, Vol. 18, pp 21-32.

10. Fach, W., Friend, J.K., Harris, R., Hickling, A., Scharpf, F., Schunk, H., and Stringer, J. 1972. Problem Structuring: a methodological approach to organisation for positive co-ordination in planning. Institute for Operational Research, report IOR/660.

11. Friend, J.K., Power, J.M., and Yewlett, C.J.L. 1974. Public Planning: the intercorporate dimension. Tavistock Publications, London.

12. Goeller, H.E. and Weinberg, A.M. 1975. The Age of Substitutability. Paper for the Science Policy Foundation 5th International Symposium.

13. Gross, B. 1967 (ed) Social Goals and Indicators for American Society. Annals of American Academy for Political and Social Science.

14. Heclo, H., and Wildavsky, A. 1974. The Private Government of Public Money. Macmillan, London.

15. Hertzberg, F. 1968. One more time: how do you motivate employees? Harvard Business Review, Vol. 46, pp 53-62.

16. Higgin, G. 1973. Symptoms of Tomorrow. Plume (Ward Lock), London.

17. HMSO 1961. Control of Public Expenditure (The Plowden Report). Cmnd 1432, London.

18. HMSO 1968. The Civil Service (Report of the Royal Commission, Chairman Lord Fulton) Cmnd 3638, London.

19. HMSO 1973. The Dispersal of Government Work from London (The Hardman Report) Cmnd 5322, London.

20. HMSO 1975. A Joint Framework for Social Policy.

21. Hutber, F.W. 1972. Review of Work on the UK National Energy Models. Department of Energy, London. Presented at ECE Symposium on Mathematical Models of the Energy Economy, USSR 1973.

22. Lindblom, C.E. 1975. The Intelligence of Democracy: decision-making through mutual adjustment. Collier-Macmillan, London. The Free Press, New York.

23. Lindblom, C.E. 1968. The Policy-Making Process. Prentice-Hall.

24. MacKenzie, W.J.M. 1969. The Plowden Report: a translation in Rose, R. Policy-making in Britain. MacMillan, London.

25. Maslow, A.H. 1954. Motivation and Personality. Harper.

26. Playford, J. 1968. The Myth of Pluralism. Arena, No. 15, pp 34-47.

27. Plessas, D.J. and Fein, F. 1972. An Evaluation of Social Indicators. Journal of the American Institute of Planners, No. 38, pp 43-51.

28. Pym, D. 1975. The Demise of Management and the Ritual of Employment. Human Relations, Vol. 28, pp 675-698.

29. Ritchie, R.S. 1971. An Institute for Research on Public Policy: a study and recommendations. Information Canada, Ottawa.

30. Scharpf, F.W. 1971. Kommplexitat als Schranke der Politischen Planung (Complexity as the limitation of political planning). German Association of Political Science, Mannheim.

31. Schon, D. 1971. Beyond the Stable State. Pelican Books, London. Random House Inc. New York.

32. Schumacher, E.F. 1973. Small is Beautiful. Blond and Briggs, London.

32. Sharpe, L.J. 1975. The Social Scientist and Policy-Making: Some cautionary thoughts and transatlantic reflections. Policy and Politics, Vol. 4, No. 2.

33. Slesser, M. 1975. Energy - the Ultimate Resource. Paper for the Science Policy Foundation 5th International Symposium.

34. Tavistock Institute and Operational Research Society 1962. An Institute for Operational Research. Report of a joint working party of the Operational Research Society and the Tavistock Institute of Human Relations.

35. Terleckyj, N.E. 1970. Measuring Progress towards Social Goals: some possibilities at national and local levels. Management Science, Vol. 16, No. 12 pp B 765-778

36. Thorsrud, E. 1968. Socio-technical approach to job design and organisational development. Management International Review, Vol. 8, pp 120-131.

37. Turner, T.P., Mills, P.M., Elton, M.C.J., Hunter, H. 1970. An Approach to the Location of Government. Paper for the TIMS International Conference, London.

38. Waddinton, C.H. 1973. OR in World War 2. Elek Science, London.

39. Willson, F.M.G. 1969. Policy-Making and the Policy Makers. In Rose, R. Policy-Making in Britain, Macmillan, London.

40. Wilson, N.A.B. 1973. On the Quality of Working Life. A report prepared for the Department of Employment, Manpower Papers No. 7, HMSO.

41. Wright, D.J. 1975. The Natural Resource Requirements of Commodities. Applied Economics, Vol. 7 pp 31-39

SUBJECT INDEX

Interdisciplinary Systems Research
Birkhäuser Verlag, Basel und Stuttgart

ISR 20
Hartmut Bossel / Salomon Klaczko / Norbert Müller (Editors):
System Theory in the Social Sciences

ISR 21
Ekkehard Brunn / Gerhard Fehl (Hrsg.):
Systemtheorie und Systemtechnik in der Raumplanung

ISR 22
Remakant Nevatia:
Computer Analysis of Scenes of 3-dimensional Curved Objects

ISR 23
Henry M. Davis:
Computer Representation of the Stereochemistry of Organic Molecules

ISR 24
Bernhelm Booss / Klaus Krickeberg (Hrsg.):
Mathematisierung der Einzelwissenschaften

ISR 25
Claus W. Gerberich
Alternativen in der Forschungs- und Entwicklungspolitik eines Unternehmens

ISR 26
Hans-Paul Schwefel:
Numerische Optimierung von Computer-Modellen mittels der Evolutionsstrategie

ISR 27
Hermann Krallmann:
Heuristische Optimierung von Simulationsmodellen mit dem Razor-Search Algorithmus

ISR 28
Stefan Rath-Nagel:
Alternative Entwicklungsmöglichkeiten der Energiewirtschaft in der BRD

ISR 29
Harry Wechsler:
Automatic Detection of Rib Contours in Chest Radiographs

ISR 30
Alfred Voss:
Ansätze zur Gesamtanalyse des Systems Mensch—Energie—Umwelt

ISR 31
Dieter Eberle:
Ein Computermodell der Verflechtung zwischen Wohn- und Naherholungsgebieten der Region Hannover

ISR 32
Ernst Billeter / Michel Cuénod / Salomon Klaczko:
Overlapping Tendencies in Operations Research, Systems Theory and Cybernetics

ISR 33
G. Matthew Bonham / Michael J. Shapiro (Editors)
Thought and Action in Foreign Policy

ISR 34
Ronald H. Atkin
Combinatorial Connectivities in Social Systems

ISR 35
Moscheh Mresse
MOSIM — ein Simulationskonzept basierend auf PL/I

ISR 36/37/38
Hartmut Bossel (Editor)
Concepts and Tools of Computer-Assisted Policy Analysis
Volume 1: Basic Concepts
Volume 2: Causal Systems Analysis
Volume 3: Cognitive Systems Analysis

ISR 39
Rolf Pfeiffer
Kybernetische Analyse ökonomischer Makromodelle für die Bundesrepublik Deutschland

ISR 40
David Canfield Smith
PYGMALION: A Computer Program to Model and Stimulate Creative Thought

ISR 41
Friedrich Niehaus
Computersimulation langfristiger Umweltbelastung durch Energieerzeugung

ISR 42
Christian König (Herausgeber)
Energiemodelle für die Bundesrepublik Deutschland

ISR 43
Werner Schülein
Anwendung des Simulationsmodells BAYMO 70 auf die Stadtentwicklungsplanung. Band 1

ISR 44
Peter Eulenberger
Anwendung des Simulationsmodells BAYMO 70 auf die Stadtentwicklungsplanung. Band 2

ISR 45
Peter Eulenberger / Werner Schülein
Anwendung des Simulationsmodells BAYMO 70 auf die Stadtentwicklungsplanung. Band 3

ISR 46
Wolfgang Birkenfeld
Methoden zur Analyse von kurzen Zeitreihen

ISR 47
Takeo Kanade
Computer recognition of human faces

ISR 48
Erwin Grochla / Wolfgang Bauer / Herbert Fuchs / Helmut Lehmann / Wolfgang Vieweg
Zeitvarianz betrieblicher Systeme

ISR 49
Derek W. Bunn / Howard Thomas
Formal Methods in Policy Formulation

ISR 50
Gisela Arndt
Planung und statistische Auswertung von Computersimulationen interdependenter Modelle mit verzögerten endogenen Variablen

ISR 51
Karl. A. Stroetmann (Editor)
Innovation, Economic Change and Technology Policies

ISR 52
Peter Sokolowsky
Grundlagen der Rechnertechnik mit einer Einführung in Mikroprozessoren

ISR 53
Claude Jablon / Jean Claude Simon
Application des modèles numériques en physique

ISR 54
Reinhard Klein
Nutzenbewertung in der Raumplanung