

OPERATIONAL TECHNIQUES FOR APPLIED DECISION
ANALYSIS UNDER UNCERTAINTY

By

Robert P. King

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Agricultural Economics

1979

ABSTRACT

OPERATIONAL TECHNIQUES FOR APPLIED DECISION ANALYSIS UNDER UNCERTAINTY

By

Robert P. King

The techniques developed in this study are designed for use during four phases of an applied decision analysis: problem formulation, the determination of subjective probability distributions, the measurement of decision maker preferences, and the identification of preferred choices. When considered together, they represent an integrated set of techniques which facilitate the application of decision theory based on the expected utility hypothesis.

Problem formulation is an important first step in any applied decision analysis. Two important considerations related to problem formulation are emphasized in this study. First, the need to identify and classify the factors which have an important impact on the outcome of the decision to be made is noted, and a classificatory scheme based on systems identification is presented. Second, the need to give careful attention to the specification of what is to be decided is stressed. The desirability of flexible decision strategies in many situations is noted, and the use of feedback control rules to introduce flexibility into a strategy is described.

Direct probability assessments of exogenous stochastic factors and the modelling of more complex stochastic processes are combined in the procedure for the determination of the distribution of outcomes associated with any choice which is presented in this study. Under

this approach, the decision maker's expectations concerning future levels of critical environmental variables are elicited directly. Monte Carlo simulation techniques are then used to determine the effect of these factors on the distribution of outcomes associated with any particular strategy. The value of this approach is greatly enhanced by the generalized procedure for the generation of sample vectors from multivariate distributions with non-normal marginals, which was developed as part of this study.

With regard to the measurement of decision maker preferences, shortcomings of both single-valued utility functions and commonly used efficiency criteria such as first and second degree stochastic dominance are identified, and a new approach to the measurement of decision maker preferences is presented. This new procedure permits the construction of interval measurements of a decision maker's absolute risk aversion.

Unlike other preference measurement function procedures, it allows the direct specification of the degree of precision with which preferences are measured, since the absolute risk aversion interval can be of any desired width. Interval measurements of this sort can be used in conjunction with the evaluative criterion of stochastic dominance with respect to a function to order alternative choices.

The final methodological contribution of this study is the formulation of a generalized risk efficient Monte Carlo programming model, which combines random search procedures, Monte Carlo simulation, and evaluation by the criterion of stochastic dominance with respect to a function within a single analytical framework for the identification

of preferred choices. This model is flexible and computationally efficient, and it is well-suited for use in the analysis of a wide range of practical decision problems.

The methodological tools developed in this study are applied to the analysis of two related problems. The first is concerned with land rental and crop production decisions on a small cash grain farm under conditions of uncertainty with respect to prices, yields, and time available for fieldwork. In the second problem analyzed, these same decisions are considered in conjunction with the selection of a flexible marketing strategy which evaluates forward contracting strategies over a seven-month period.

For Jane

ACKNOWLEDGEMENTS

For their guidance and assistance, I wish to thank the members of my thesis committee: Roy Black, Ralph Hepp, Glenn Johnson, and Lindon Robison. I am especially grateful to Lindon Robison, my thesis supervisor, whose questions prompted this study and whose willingness to explore new ideas was a continuing source of help. I also want to give particular thanks to Warren Vincent, who was my major professor throughout most of my graduate career. Without the freedom and the good advice he gave me this study might not have been undertaken.

Funding for this study was provided by the Michigan Agricultural Experiment Station. I gratefully acknowledge their support.

Finally, I thank my wife, Jane, who shared all the good and bad days that come with an undertaking like this one. Her encouragement and sense of perspective were a constant source of help.

TABLE OF CONTENTS

| | |
|---|-------------|
| LIST OF TABLES | vi |
| LIST OF FIGURES | viii |
| <u>Chapter</u> | <u>Page</u> |
| I. INTRODUCTION | |
| 1.1 Background and Need for the Study | 1 |
| 1.2 Problem Statement | 4 |
| 1.3 Objectives of the Study | 8 |
| 1.4 Plan for the Remainder of the Study | 9 |
| II. PROBLEM FORMULATION IN THE ANALYSIS OF DECISIONS MADE UNDER UNCERTAINTY | |
| 2.1 Introduction | 12 |
| 2.2 The Identification and Classification of Factors Relevant to the Decision | 14 |
| 2.3 The Specification of the Decision to be Made | 21 |
| 2.3.1 The Specification of Feedback Control Rules | 23 |
| 2.3.2 Plans, Decisions, and Actions | 28 |
| 2.4 A Formal Statement of the Decision Problem | 33 |
| 2.5 An Application | 35 |
| III. THE DETERMINATION OF SUBJECTIVE PROBABILITIES | |
| 3.1 Introduction | 44 |
| 3.2 Probability Encoding Procedures | 47 |
| 3.3 Modelling Stochastic Processes | 56 |
| 3.4 An Application | 65 |
| IV. THE MEASUREMENT OF DECISION MAKER PREFERENCES | |
| 4.1 Introduction | 84 |
| 4.2 The Use of Single Valued Utility Functions to Represent Decision Maker Preferences | 87 |
| 4.3 Efficiency Criteria and the Representation of Decision Maker Preferences | 91 |
| 4.4 Stochastic Dominance with Respect to a Function | 96 |
| 4.5 An Interval Approach to the Measurement of Decision Maker Preferences | 105 |
| 4.6 Implementation of the Procedure | 109 |
| 4.7 An Empirical Test | 121 |
| 4.8 An Application | 125 |

| <u>Chapter</u> | <u>Page</u> |
|--|-------------|
| V. COMPUTATIONAL PROCEDURES FOR THE IDENTIFICATION OF PREFERRED CHOICES | |
| 5.1 Introduction | 130 |
| 5.2 A Review of Existing Computational Procedures . . . | 131 |
| 5.3 A Generalized Procedure for the Identification . . . | 137 |
| of Preferred Choices Under Uncertainty | |
| 5.3.1 Generation of a Feasible Management Strategy. | 139 |
| 5.3.2 Determination of the Distribution of | 146 |
| System Output Levels | |
| 5.3.3 The Evaluation of Alternative Strategies . . | 147 |
| 5.3.4 General Comments on the GREMP Model | 151 |
| 5.4 An Application | 156 |
| VI. COMBINED PRODUCTION AND MARKETING DECISIONS BY CASH GRAIN FARMERS: AN EXTENDED APPLICATION | |
| 6.1 Introduction | 162 |
| 6.2 Problem Formulation | 166 |
| 6.3 The Determination of Subjective | 172 |
| Probability Distributions | |
| 6.4 Decision Maker Preferences | 179 |
| 6.5 The Identification of Preferred Choices | 180 |
| 6.6 Further Discussion of the Results | 191 |
| 6.7 Implications for Further Research | 194 |
| VII. SUMMARY AND CONCLUSIONS | |
| 7.1 A Review of the Methodological Tools | 197 |
| Developed in this Study | |
| 7.2 Empirical Findings | 201 |
| 7.3 Implications for Future Research | 203 |
| APPENDIX A | 207 |
| A GENERALIZED MULTIVARIATE PROCESS GENERATOR | |
| APPENDIX B | 240 |
| IMPLEMENTATION OF THE INTERVAL APPROACH TO THE MEASUREMENT OF DECISION MAKER PREFERENCES | |
| APPENDIX C | 278 |
| IMPLEMENTATION OF THE GREMP MODEL | |

LIST OF TABLES

| <u>Table</u> | <u>Page</u> |
|--|-------------|
| 2.1 Standard Crop Enterprise Budgets | 37 |
| 2.2 Planting and Harvest Periods and Possible Crops for Each Combination | 39 |
| 3.1 Average Corn Yield and Moisture Content by Planting and Harvest Date | 67 |
| 3.2 Average Soybean Yield by Planting and Harvest Date . . . | 68 |
| 3.3 Percent of Time Available for Fieldwork by Calendar Period for Well Drained Sandy Loom Soils in the Lenawee, Monroe, Livingston County Area | 71 |
| 3.4 Beta Distribution Parameters for Time Available for Fieldwork by Calendar Period | 72 |
| 3.5 Beta Distribution Parameters for Corn and Soybean Price Distributions | 74 |
| 3.6 Two Possible Management Strategies | 78 |
| 3.7 System Performance under Strategy 1 | 79 |
| 3.8 System Performance under Strategy 2 | 81 |
| 4.1 Probability Distribution Associated with Two Alternative Action Strategies | 90 |
| 4.2 Sample Distribution from a Normal Distribution With $\mu = 3000$ and $\sigma = 1000$ | 115 |
| 4.3 Boundary Intervals for Pairs of Sample Distributions | 116 |
| 4.4 Performance Indicators for Alternative Preference Measures | 124 |
| 5.1 Efficient Strategies for Decision Maker A | 158 |
| 5.2 Efficient Strategies for Decision Maker B | 159 |
| 5.3 Efficient Strategies for Decision Maker C | 161 |
| 6.1 System Performance under a Sample Management Strategy | 178 |
| 6.2 Representative Strategies from the Efficient Sets of Four Decision Makers | 182 |
| 6.3 System Performance under the Preferred Strategy of . . . a Risk Neutral Decision Maker | 183 |
| 6.4 System Performance under the Representative Strategy for Decision Maker A | 186 |
| 6.5 System Performance under the Representative Strategy for Decision Maker B | 188 |
| 6.6 System Performance under the Representative Strategy for Decision Maker C | 190 |

| <u>Table</u> | <u>Page</u> |
|---|-------------|
| A.1 Parameters of Marginal Beta Distributions | 236 |
| A.2 Specified Correlation Matrix | 236 |
| A.3 Sample Correlation Matrix for Multivariate Normal Data | 238 |
| A.4 Sample Correlation Matrix for Multivariate Beta Data | 238 |
| B.1 An Example of Preference Data Input for Program UFUNC | 270 |

LIST OF FIGURES

| <u>Figure</u> | <u>Page</u> |
|--|-------------|
| 2.1 A General Scheme for System Identification | 16 |
| 3.1 A Cumulative Distribution Function | 51 |
| 3.2 A Cumulative Distribution Function Based on Five Sample Observations | 62 |
| 3.3 Cumulative Distribution Functions for Two Alternative Strategies | 82 |
| 4.1 Illustrations of First and Second Degree Stochastic Dominance | 93 |
| 4.2 Cumulative Distribution Functions $F(y)$ and $G(y)$ | 102 |
| 4.3 A Graph of the Function $[G(y)-F(y)]$ | 103 |
| 4.4 A Sequence of Interval Preference Measurements | 110 |
| 4.5 Upper and Lower Bound Absolute Risk Aversion Functions Based on Three Interval Measurements | 111 |
| 4.6 An Absolute Risk Aversion Measurement Scale | 113 |
| 4.7 A Sample Questionnaire | 119 |
| 4.8 Interval Measurements of Absolute Risk Aversion for Three Decision Makers | 128 |
| 5.1 A Flow Chart of the GREMP Model | 140 |
| 5.2 Sample Observations from the Distribution of Outcomes Associated with Strategy One | 148 |
| 5.3 Sample Observations from the Distribution of Outcomes Associated with Strategy Two | 150 |
| 5.4 Sample Observations from the Distribution of Outcomes Associated with Strategy Three | 152 |
| 6.1 Dates for the Application of Forward Contracting Rules | 173 |
| 6.2 Price Forecast Line and Observed Control Price Levels for One State of Nature | 175 |
| 6.3 The Expected Net Return Maximizing Strategy | 193 |

| <u>Figure</u> | <u>Page</u> |
|---|-------------|
| A.1 Generation of Random Variates by the Inverse Transformation Method | 212 |
| A.2 A Program for Generating Exponential Random Variates | 214 |
| A.3 A Program for Generating Gamma Random Variates | 216 |
| A.4 A Program for Generating Beta Random Variates | 219 |
| A.5 An Approximate Representation of a Cumulative Distribution Function Based on Six Known Points | 221 |
| A.6 A Table Look-up Function Subprogram | 222 |
| A.7 A Generalized Univariate Process Generator | 224 |
| A.8 A Program for Generating Multivariate Normal Random Variates | 227 |
| A.9 A Generalized Multivariate Process Generator | 231 |
| A.10 A Multivariate Beta Process Generator | 234 |
| B.1 Sequence of Choices for a Measurement of Absolute Risk Aversion Based on Three Questions | 242 |
| B.2 A Suggested Absolute Risk Aversion Measurement Scale | 244 |
| B.3 A Listing of Program NORGEN | 245 |
| B.4 Sample Output from Program NORGEN | 248 |
| B.5 A Listing of Program INTID | 250 |
| B.6 An Eight Element Measurement Scale | 255 |
| B.7 Sample Output for Program INTID | 256 |
| B.8 A Three-Stage Hierarchy of Questions | 258 |
| B.9 Sample Questionnaire | 261 |
| B.10 A Listing of Program UFUNC | 264 |
| B.11 Representation of a Utility Function by Interpolation Between Known Points | 266 |
| B.12 An Interval Preference Measurement | 268 |
| B.13 Utility Functions Associated with Upper and Lower Bound Absolute Risk Aversion Functions | 271 |
| B.14 A Listing of Program NSTDO | 272 |
| B.15 A Sample Output from Program NSTDO | 276 |
| C.1 General Flow Chart of Program GREMP | 279 |
| C.2 Statements in GREMP Which Generate Feasible Strategies | 286 |
| C.3 A Simple Version of Subroutine CHECK | 288 |
| C.4 A Simple Version of Subroutine DISGEN | 291 |

CHAPTER I

INTRODUCTION

1.1 Background and Need for the Study

Uncertainty is that state of knowledge in which the consequences of actions being considered cannot be specified exactly; it is that condition in which knowledge is to some degree imperfect or incomplete.¹ As such, the presence of uncertainty is a basic fact in nearly all decision situations.

Though almost always present to some degree, uncertainty is not a factor which must be considered explicitly in every instance. Often the consequences of ignoring imperfections in knowledge are judged to be minimal, and decisions can be made as though all relevant factors were known with certainty. In many other instances, however, when the outcome of an important choice is highly uncertain, the failure to consider uncertainty explicitly may not be justifiable.

Uncertainty can have an important impact both on the process by which decisions are made and on the character of decisions themselves. Learning, which has no value when knowledge is perfect, becomes a

¹ Knight's (1921) distinction between risk and uncertainty is not made in this study. General acceptance of Ramsey's (1931) observation that knowledge of the true probability distribution of a random variable is not possible obviates the need for this distinction. It is important to recognize, however, that degrees of uncertainty can vary and that a decision maker's state of knowledge has an important impact on his actions (Wald, 1947; Johnson and Lard, 1961).

potentially worthwhile activity under uncertainty. As a result, the decision maker's attention may be focused primarily on the decision of whether or not to continue learning rather than on the actual choice of an action to be undertaken.¹ When learning stops and a choice of actions is made, the character of that choice may also differ from that of one made under certainty. For example, decisions made under uncertainty often take the form of flexible strategies which make forthcoming actions contingent upon future events that the decision maker can observe but cannot control.² Such strategies would be of little value if knowledge were perfect and all future occurrences could be known in advance. The presence of uncertainty also affects the character of the decision rules used to identify a preferred choice. Decision rules which give explicit consideration to uncertainty generally require more specific information about decision maker preferences than is required under certainty, and they must permit the synthesis of this normative information with probabilistic information about the possible outcomes of any choice being considered.

Despite these and other impacts of uncertainty in the decision process, uncertainty is often not considered explicitly in the analysis of decisions upon which it may have a profound effect. In many instances the failure to give proper attention to such an important factor is not due to a lack of recognition of the impacts of uncertainty or to

¹The role of learning in the decision process is discussed in Bradford and Johnson (1953) and in Johnson and Lard (1961). More recently, the work of Aoki (1975), MacRae (1975) and others have dealt with the "dual-control" problem of learning and setting policies simultaneously.

²Massé (1962), Cocks (1968), Rae (1971), and Day (1975) have all stressed the importance of such adaptive decision strategies.

inadequacies in the theory of decision making under uncertainty. Rather, this failure can often be attributed to a lack of workable analytical techniques which permit the explicit consideration of uncertainty in the analysis of practical decision problems, techniques which are flexible enough to allow the application of powerful theoretical results in a wide range of complex situations.

This study is concerned with the development of an integrated set of techniques which facilitate the incorporation of explicit considerations of uncertainty into a decision analysis. Particular emphasis is placed on the development of methodological tools which make the application of decision theory based on the expected utility hypothesis more feasible in a wide range of practical problem solving contexts. The expected utility hypothesis has been the basis for much of the body of theory concerned with decision making under uncertainty and has been used to explain a diverse range of behavioral patterns.¹ It is also a potentially powerful tool for the analysis of decision problems in a practical context. But for a few notable exceptions such as Grayson (1960), Howard, Matheson, and North (1972), and Keeney (1973), however, this body of theory has rarely been applied successfully in the solution of practical decision problems.

A number of difficulties have limited the usefulness of decision theory based on the expected utility hypothesis. The expectations and preferences of decision makers have proved to be difficult to determine

¹Notable among the theoretical applications of the expected utility hypothesis are the important early articles by Friedman and Savage (1948), the more recent work of Samuelson (1967) and Ehrlich and Becker (1972), and the extensive literature concerned with portfolio selection they based on the work of Tobin (1958), and Markowitz (1959) and Baumol (1970), among others.

and represent accurately, and, as Zadeh (1973) and Watson, Weiss, and Donnell (1979) note, it is somewhat paradoxical that extremely precise decision rules based on the expected utility hypothesis are applied in situations where relevant information is highly imprecise. Computational problems associated with the implementation of expected utility maximizing decision rules are also a source of difficulty. Often they force the imposition of restrictive assumptions on the way expectations and preferences are represented and so further limit the theory's usefulness in an applied context (Anderson, 1975). Finally, as Johnson (1976), Day (1964), Cyert and March (1963), and others have observed, much more than the appreciation of a decision rule is involved in the choice process. Even when other methodological difficulties can be resolved, decision theory based on the expected utility hypothesis can be successfully applied only if a better understanding of problems and the process by which they are resolved is attained.

These difficulties are serious ones, but they stem from problems with the way decision theory has been applied rather than from the theory itself. They point to a need to develop methodological tools which can be used to make decision theory based on the expected utility hypothesis truly operational in a practical context.

1.2 Problem Statement

In response to this need, this study focuses on the problem of formulating an integrated set of operational techniques for the analysis of decision making under uncertainty, techniques which are consistent with theory based on the expected utility hypothesis and which overcome a number of the problems encountered in earlier attempts to apply that

theory. Within this broad problem, four more specific areas of difficulty can be identified: problem formulation, the determination and representation of expectations, the measurement of decision maker preferences, and the identification of preferred choices.

A problem is said to exist when "a condition, situation, or thing is not as good or is worse (more bad) than it could be," (Johnson, 1976, p. 270). Before the information needed to resolve a problem can be collected and analyzed, before a course of action can be selected and implemented, the problem itself must be clearly defined. Problem formulation is the process by which such a problem definition is developed. It requires that performance criteria, choice variables, and relevant factors in the decision situation which are beyond the control of the decision maker must be defined and that the nature of the decision to be made must be clearly specified. Despite its importance, problem formulation is often given relatively little attention. Frequently, for example, the definition of a problem under consideration is dictated by the computational tools to be used as aids in its resolution. As a result, important sources of uncertainty may not be considered and the special character of decisions made under uncertainty may be ignored. Despite insights provided by Johnson (1961a), Cyert and March (1963), Churchman (1968), Day (1971, 1975) and others, then, problem formulation continues to be a problem in the analysis of decisions made under uncertainty.

The process of selecting a course of action which will best resolve a particular problem requires the synthesis of two types of knowledge: (1) positive knowledge, which pertains to beliefs about what is, what

will be, and what can be done; and (2) normative knowledge, which pertains to beliefs about the goodness or badness of particular conditions, situations, and things. The product of such a synthesis is prescriptive knowledge, which allows the decision maker to prescribe or specify the right strategy or set of actions.¹

Expectations concerning the relative likelihood of alternative events occurring in the future represent an important part of the positive information required in any applied decision analysis. A number of methodological problems arise in connection with the determination and representation of expectations. Frequently, they are not well formulated in the mind of the decision maker, who may not be accustomed to thinking in probabilistic terms or may simply not be familiar with a particular factor which may have a significant impact on the consequences of his choice. Even when the decision maker's expectations are well formulated, problems may arise because he is unable to express them in a form which is useful analytically. Other difficulties may stem from the fact that the process by which the outcomes of particular actions are determined may be so complex that it cannot be comprehended as a unified whole. Given these problems, there is a need for procedures which help the decision maker structure his own thoughts and help him to use information from more expert sources, a need for techniques which allow the decision maker to break down complex processes

¹The distinctions between "right" and "wrong" and "good" and "bad" made by Lewis (1955), is an important one. The adjectives "right" and "wrong" refer to the nature of an act, while "good" and "bad" refer to the conditions existing prior to an act or to its consequences. To say a condition or consequence is good is the statement of a normative belief. To say that an act is right is the statement of a prescriptive belief which is founded both on positive information concerning the consequences of the act and on normative information pertaining to the quality of goodness of the consequences.

into more comprehensible sub-processes about which expectations can be more easily formulated and then reintegrate that information for use in the decision analysis.

Information on decision maker preferences is the primary normative input in any decision analysis. Problems associated with the measurement and representation of preferences also cause serious difficulties in the analysis of decisions made under uncertainty. Currently available measurement techniques are used to construct single valued utility functions, which are precise but often inaccurate representations of preference, and many place little faith in them. Efficiency criteria based on stochastic dominance, on the other hand, require little specific information about the decision maker's preferences, but they often fail to order choices and may not eliminate enough alternatives when a large number must be evaluated. These difficulties indicate that there is a need for preference measurement techniques which are more reliable and easier to use in an applied context.

The fourth major area of difficulty is that of identifying a preferred choice or set of choices from what may be an infinitely large number of alternatives. This requires the synthesis of both positive and normative information, the simultaneous consideration of both expectations and preferences. Mathematical programming techniques are commonly used as computational aids in the solution of complex quantitative problems. They are best suited, however, for use in situations where uncertainty is not a major factor. As Anderson (1975) notes, the use of mathematical programming in the analysis of decisions made under uncertainty often requires that rather severe restrictions be placed on the manner in which a decision problem is posed and in the way

information on expectations and preferences is represented. These difficulties point to the need for more flexible computational procedures.

1.3 Objectives of the Study

The problems which motivate this study are primarily methodological. They have important implications, however, for the analysis of decision problems in a practical context. In response to them, the objectives of the study are:

1. To present a framework for problem formulation which can serve as a guide in the identification and structuring of information required for the analysis of decisions made under uncertainty.
2. To review procedures used in the elicitation of information on decision maker preferences and to refine techniques for the determination of probability distribution for outcomes which are the result of complex processes affected by a variety of stochastic factors.
3. To develop and test a technique for the measurement of decision maker preferences which is well suited for use in an applied context and which overcomes some of the difficulties associated with other preference measurement procedures.
4. To formulate and make operational a computational procedure for the identification of preferred choices which is flexible enough to be used in the analysis of a wide range of practical problems and which imposes few restrictions on problem formulation or on the representation of decision maker expectations and preferences.

The fulfillment of each of these four objectives contributes to the primary purpose of this study, which is to develop an integrated set of techniques for use in the analysis of decisions made under uncertainty. Emphasis should be placed on the fact that the usefulness of the techniques presented below is greatly enhanced by their having been combined within a single methodological framework.

1.4 Plan for the Remainder of the Study

The principle objectives of the study are addressed in the next four chapters. Problem formulation is the subject of Chapter II. Particular emphasis is given to the usefulness of system identification (Manetsch and Park, 1977a) as an aid in structuring information in a practical decision context. The need to recognize the importance of opportunities for learning and adaptive behavior and the impact such opportunities have on the character of decisions is also stressed.

Techniques for determining decision maker expectations are described in Chapter III. Procedures for eliciting information on subjective probability distributions are reviewed, and the use of Monte Carlo simulation techniques to model the performance of complex stochastic systems for which outcome distributions cannot be determined analytically is discussed. The value of this approach is greatly enhanced by the generalized multivariate process generator developed as part of this study. Described in detail in Appendix A, this analytical tool can be used to model multivariate probability distributions defined by marginals of any form and by any positive definite correlation matrix.

The measurement of decision maker preferences is the subject of Chapter IV.¹ Procedures used to derive single valued utility functions are first reviewed, and several commonly used efficiency criteria are discussed as possible alternatives to the use of single valued utility functions in applied decision analyses.² A more general and more powerful efficiency criterion, stochastic dominance with respect to a function (Meyer, 1977a), is then described and a new approach to the measurement of decision maker preferences, developed as part of this study for use in conjunction with this criterion, is introduced. This new approach allows the analyst to specify the degree of precision with which decision maker preferences are measured. Results of an experimental test of this technique are also presented. They demonstrate its flexibility and its predictive power.

Computational procedures for the identification of preferred choices are the subject of Chapter V. Uses of mathematical programming techniques in the analysis of decisions made under uncertainty are reviewed first, and the major shortcomings of these techniques are identified. A new procedure for the identification of preferred choices which combines random search methods, simulation techniques, and evaluation by the criterion of stochastic dominance with respect to a function

¹Though the importance of multiple objectives in many decision situations is recognized, attention in this study focuses entirely on preference relationships which depend only on the level of a single performance criterion.

²Computer programs used in the implementation of this technique are listed in Appendix B.

is introduced and described in detail.¹ This computational tool is remarkably flexible, placing few restrictions on the way probabilities and preferences are represented or on the general form of the problem to be solved.

Two related examples are used to illustrate the techniques developed in this study. Both are concerned with decisions affecting the operation of a southeastern Michigan cash grain farm. The simpler of the two examples, which focuses on land rental and crop mix decisions under price and yield uncertainty, is discussed at the end of the four methodological chapters. It is used to demonstrate how techniques for problem formulation, the determination of expectations, the measurement of preferences, and the identification of preferred choices can actually be applied. The second example, which considers the selection of a marketing strategy in conjunction with production and land rental decisions, is the subject of Chapter VI. Again prices and yields are uncertain. The marketing strategies considered may include cash sales at harvest, forward contracting, or any combination of these. Emphasis is placed on the adaptive nature of such strategies and on the impact of preferences on the combined production-marketing strategy selected.

Finally, in Chapter VII the strengths and weaknesses of the integrated set of techniques developed in this study will be discussed. Particular attention will be given to an evaluation of the range of applications for which these techniques can be of use and to the identification of areas where further methodological improvements are needed.

¹The computer program which implements this decision model is presented in Appendix C.

CHAPTER II

PROBLEM FORMULATION IN THE ANALYSIS OF DECISIONS MADE UNDER UNCERTAINTY

2.1 Introduction

A problem exists "when an indeterminate situation, present or projected, is regarded as unsatisfactory and a more satisfactory alternative situation is sought" (Johnson and Zerby, 1973, p. 3). Management is the process by which problems of a practical nature are resolved.¹ In describing the managerial process, Johnson (1976) has identified six major types of activities: problem definition, observation, analysis, decision, execution, and responsibility bearing. This study is concerned primarily with the development of analytical tools which can aid the decision maker during the analysis and decision phases of the management process--tools which facilitate the determination of distributions of outcomes associated with alternative actions, the measurement of decision maker preferences, and the application of decision rules used to identify preferred choices. These tools can be of little use in a practical context, however, if the problems to which they are applied have not been clearly and correctly specified.

¹As Johnson and Zerby (1973) note, problems can be practical or theoretical in nature. Practical problems are those which are related to the choice of an action and so demand some form of resolution. Theoretical problems, on the other hand, are not linked to a definite action which can be fixed in space and time and so may never be fully resolved.

Problem formulation is the process by which a problem is defined and structured for analysis. It is the process by which features of the problematic situation judged to have an important impact on the choice to be made are identified and classified and the process by which the nature of the decision to be made is specified. The product of this process should serve as a guide in the collection of additional information, should provide a framework for the organization of that information, and should help structure the analysis which leads to a decision. As such, problem formulation is a critical activity within any applied decision analysis. Major expenditures of resources for information gathering and analysis may be required before a choice can be made in some decision situations. These resources can be used effectively only if the problem under consideration is clearly and correctly defined. A carefully determined solution to an irrelevant problem is of little use.

Problem formulation is the subject of this chapter. The discussion in subsequent sections focuses on two important aspects of the process of problem formulation: the identification and classification of variables relevant to the analysis of a particular problem and the actual specification of what is to be decided. With respect to both of these activities, emphasis is placed on the need to recognize the dynamic character of many decision situations and the impact it has on the choice process. The need to recognize the role of learning and the effect it has both on the way problems are formulated and the manner in which decisions are made is also stressed. The purpose of this discussion is not to introduce new concepts or to develop a comprehensive procedure for problem formulation. Rather, it is to restate some

valuable observations made by others and to present a general view of decision problems and a working vocabulary for the discussion of them that can provide insights into the process of problem formulation.

2.2 The Identification and Classification of Factors Relevant to the Decision

At the outset of a decision analysis the problem under consideration may be only vaguely defined in the mind of the decision maker. To better understand the nature of the problem, one of the first tasks usually undertaken is the identification of factors judged to have an important impact on the choice to be made. Not only does this help to clarify the problem, but it also establishes a set of variables which can be the focus of observations and analysis. As more is learned about the problem, this set of relevant factors is, of course, repeatedly revised.

Efforts to identify the important factors in a particular decision situation are facilitated by the presence of a general classificatory framework which suggests broad types of variables that should be considered. Such a framework is presented by Manetsch and Park (1977a) in their discussion of system identification. A system can be defined as a collection of objects or processes which interact to perform a given function or set of functions. System identification is a generalized scheme for structuring information about the characteristics of particular systems. It is a particularly valuable classificatory framework in a practical problem solving context because it is well suited for the description of static as well as dynamic decision problems and because it encourages the explicit consideration of the sources of uncertainty in any particular situation.

Manetsch and Park identify five broad classes of variables which should be considered in any decision situation: system outputs, controllable system inputs, exogenous system inputs, system state variables, and system design parameters. These are the major elements in Figure 2.1. Before discussing each of these categories it should be noted that a "system" is a concept rather than an actual entity. The definition of a system--the specification of its functions and component processes--depends on the purposes for which it is being considered. In the analysis of a particular decision, the function of the system considered should be to determine or at least affect the situation or condition which is judged to be problematic, and the system's component processes should include all those which have an important impact on that situation. For example, the problem facing a farm family may be that its standard of living is unacceptably low. In trying to improve this unsatisfactory situation they will want to consider the system whose functions it is to provide them with the resources for obtaining food, clothing, shelter, and other necessities. The component processes within that system would be that set of processes by which such resources are generated--a set which might include farm production and marketing, off-farm work, and public assistance.

System identification begins with the specification of system output variables. System outputs are the products of the processes which comprise a particular system; and system output variables, which measure levels of system outputs, should serve as indicators of the degree to which the system under consideration performs its designated functions. They should be the basis for a reliable representation of

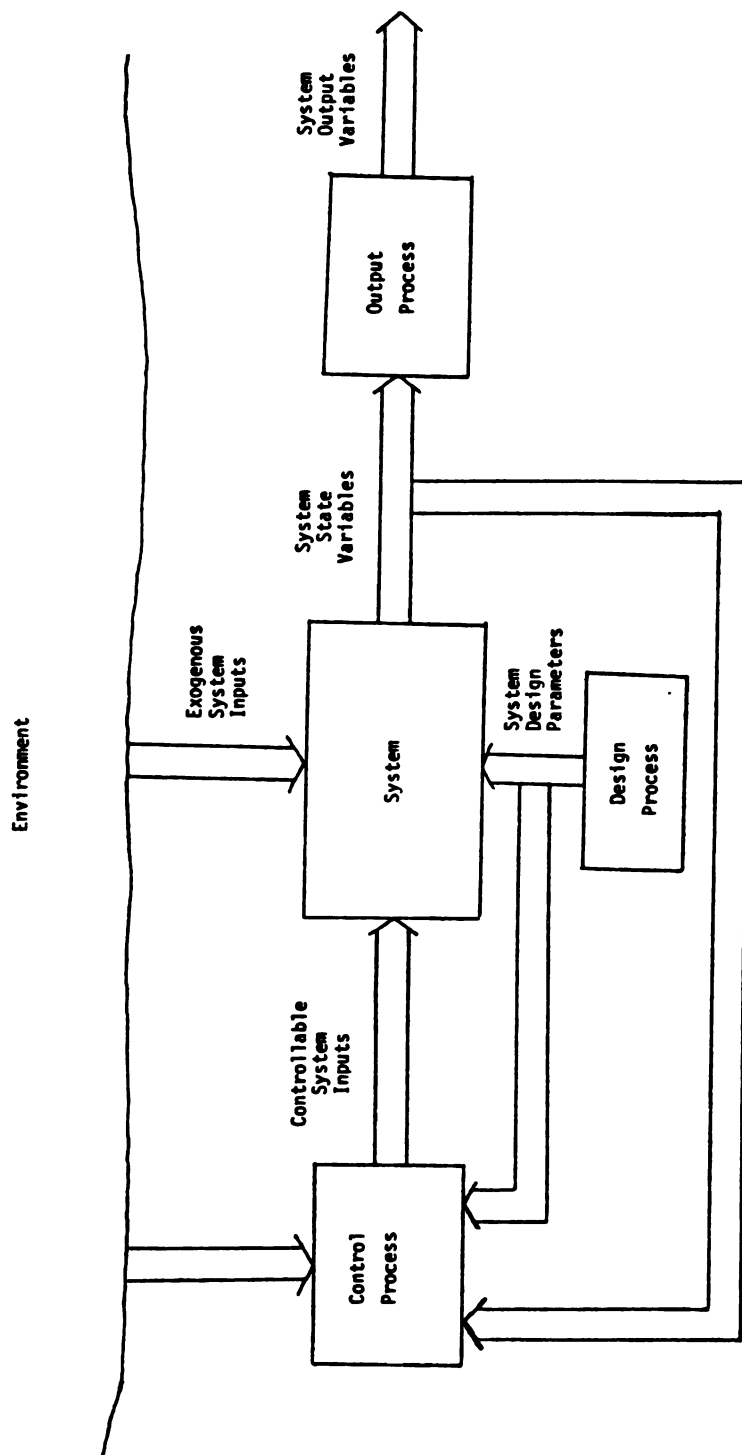


Figure 2.1 A General Scheme for System Identification

all relevant features of system performance, providing information about both the desired and undesired impacts of any choice being considered. For this reason, considerable care should be taken in identifying system output variables. In some situations all relevant information may be conveyed by a single output measure, but in other instances more than one system output variable is needed to adequately represent system performance. For the family in our example, net annual household income may be an adequate measure of system performance; but it may be necessary to specify other system output variables if, for example, there are important costs associated with the acceptance of public assistance which are not reflected by the level of net annual income.

System output levels are determined by inputs to the system and by its structure. System inputs are factors or stimuli emanating from outside the system which affect its performance. They can be classified as controllable or exogenous. Controllable system inputs are those for which a level can, to some degree, be specified by the decision maker. The level of a controllable system input may represent an amount of some physical factor of production flowing into one of the processes within the system or it may specify a level of some well defined activity. In our example the set of controllable system input variables might include designations of levels for each farm production enterprise and for hours of off-farm work activity and a binary variable indicating participation or non-participation in public assistance programs.

The levels of exogenous system inputs cannot be determined by the decision maker. Rather, they are determined by the system environment, a set of processes which affect system performance but are not, in turn,

significantly affected by the system's behavior. The set of exogenous system inputs in our example might include farm product prices, levels of rainfall, wage rates, and levels of public assistance. All of these have a potentially important impact on the family's standard of living but are beyond its control. The distinction between the system and its environment is not always evident, nor is it necessarily fixed. It depends on the problem under consideration and on the power of the decision maker. The distinction is an important one, however, especially in the analysis of decisions made under uncertainty, since stochastic factors in the environment can be viewed as the primary source of uncertainty in most decision situations.

System structure determines the relationship between system inputs and system outputs. The structure of a system is described by system state variables and by system design parameters. State variables are descriptors of the state or condition of a system at any point in time. In general system outputs can be viewed as functions of the system's state at some specified time or as functions of the system's state through time. In addition to determining system output levels, the state of the system may also affect the range of allowable levels for controllable system inputs. Therefore, it is important to give careful consideration to the specification of system state variables. In our example the set of system state variables related to crop production processes might include current levels of acreage planted to each crop grown, current amount of each crop harvested, current crop production expenses incurred, and current crop sales receipts.

System design parameters define the relationship between inputs to the system and its resultant state. As such they describe the processes

which comprise the system. With regard to the crop production processes in our example, the set of system design parameters could include variable production costs per acre of each crop grown, time required to plant or harvest an acre of each crop grown, and parameters indicating tillage practices and standard procedures concerning the order of procedures for the planting and harvesting of each crop. System design parameters have an important impact on system performance, and in some instances they can be altered by the decision maker. For example, a change in tillage practices may significantly affect both timeliness and crop yields for the farm in our example. Such a change, however, may be costly. As a result, alterations in system design are usually undertaken only in response to serious problems which cannot be resolved by other means.

In any decision analysis variables in each of these categories should be identified. The analysis itself focuses on the specification of a strategy to be undertaken to resolve the problem being considered. A strategy is defined by desired levels for controllable system input variables and by the new values of any system design parameters that are to be changed. Manetsch and Park (1977a) define "management" or "control" as the process by which desired levels for controllable system input levels and "design" or "planning" as the process by which specifications for system structure are made.¹ In general there are limits on the range of strategies that can be undertaken in any particular situation. As indicated earlier in Figure 2.1, the state of the system,

¹This concept of management is a much more narrow one than that used in this study. The management of a system is but one activity within the broader managerial process.

system design parameters, and factors in the environment all affect the control process. They impose constraints which restrict the allowable range of values for controllable system input variables. Though not shown explicitly in Figure 2.1, similar constraints also limit design changes. As the process of system identification proceeds, then, it is also important to identify those factors which restrict the range of available choices.

To this point in the discussion systems have been viewed as essentially unidirectional processes which convert inputs to outputs. This view is limited because it fails to recognize the fact that as strategies are implemented there may be opportunities for learning and for revising chosen plans of action on the basis of newly acquired information. Introduction of the concept of feedback into our view of system identification helps to overcome this limitation. Feedback is the return flow of information (both positive and normative) on the state of the system and into environment to the central process unit. Recognition of the feedback loop between the set of system state variables and the control unit changes our conceptualization of the process by which inputs are transformed to outputs from one which is essentially unidirectional and disjointed to one in which this process is viewed as a continuous cycle or closed loop. This is a more realistic way of representing the context in which decisions are made, especially in situations where the impact of uncertainty is important. Most decisions are not made in isolation, nor are they implemented instantaneously. Rather, they are made sequentially, and the outcome of one decision affects the opportunity set from which future choices can be made. Furthermore, because decisions are implemented over a

period of time, there are often opportunities to revise them. In such a context, learning is of considerable importance, and feedback in the medium through which learning takes place. The recognition of such informational flows can have an important impact on the specification of what is to be decided. Therefore, there is a need to consider such factors during the process of problem formulation.

Finally, it should be noted that the identification and classification of the important factors in a particular decision situation is, in itself, a learning process. A decision maker's view of a problem and of the set of processes by which that problem can be resolved is repeatedly revised and made more complete. This kind of learning continues until further efforts are judged not to be worthwhile or until the decision maker is forced to take an action.¹ The degree of detail included in the description of a particular system, then, depends on the usefulness of that detail in helping the decision maker determine his preferred course of action.

2 - 3 The Specification of the Decision to be Made

As was noted in Chapter I, the presence of uncertainty can have an important impact on the character of decisions. All situations involving uncertainty are, in a sense, dynamic, since they are characterized by changes in the decision maker's knowledge through time. One makes a choice and begins to act in the present, but only later do the consequences of one's actions come to be known. Often the outcomes

¹Johnson and Lard (1961) relate the decision of whether or not to continue learning to five more formally defined knowledge situations: learning, forced learning, forced action, inaction, and risk.

associated with a particular strategy unfold over an extended period of time, and there are opportunities for learning as the strategy is implemented. The existence of such opportunities may make it desirable to introduce flexibility into the specification of a decision strategy. The strategy becomes a conditional plan--a set of contingency rules which direct actions on the basis of currently available information.¹ In this way the importance of future opportunities for learning is recognized explicitly when a choice is made. Though not all decisions made under uncertainty take this form, many do. The discussion in this section will focus, in part, on the specification of flexible strategies based on feedback control rules.

Another important consideration which affects the character of decisions is the length of the planning horizon. Because current choices have an impact on future opportunities, it is often necessary to formulate strategies which extend into the future. When knowledge is perfect, it is possible to specify future actions extending over an infinite planning horizon. When knowledge is not perfect and reliable information about future events can be attained only at considerable cost, on the other hand, the time horizon for which it is worthwhile to formulate a plan may be shortened considerably (Modigliani and Cohen, 1961; Kleindorfer and Kunreuther, 1978). Specification of the relevant planning horizon and the distinctions among a plan, a decision, and an action, then, will also be considered in this section.

¹Dreyfus (1968) demonstrates that flexible strategies are superior to inflexible ones in multistage decision problems and uncertainty.

Before beginning the discussion it should be noted that the specification of what is to be decided, like the identification and classification of the important factors in a decision situation, is a process which continues throughout a decision analysis. As more is learned about the problem at hand, as observation and analysis continue, one's conceptualization of what is to be decided is repeatedly revised, and refinements in this aspect of problem formulation continue until further changes are not worthwhile.

2.3.1 The Specification of Feedback Control Rules

A feedback control rule is a rule for processing current information on the state of a system and its environment in order to repeatedly update desired controllable system input levels. Feedback control rules can take a variety of forms. They can be as simple as the statement, "If the forward contract price of corn is below \$2.00 on May 1, I'll comply with the federal set aside program requirements; if it's above \$2.00, I won't participate." Alternatively, they can be complex functions of several state variables. In specifying a feedback control rule to determine the level of some controllable system input variable at any point in time, one must be concerned with the identification of state variables which can be expected to have a significant impact on the desired level of the controllable system input being considered, with the form of the rule, and with the actual parameters of the rule.

The simple feedback control rule stated above determines the level of a binary controllable system input variable which has a value of 0 if the operator chooses not to participate and 1 if he chooses to participate. The only state variable affecting the choice of a level for

this variable is the forward contract price of corn. The form of the rule is that of an "if-then" statement, and the parameters of the rule are a forward contract price of corn--\$2.00--and a date when a decision will be made--May 1. It should be noted that the effectiveness of the rule is affected by the variables considered, by its form, and by its parameters. All affect its impact on system performance, and in selecting a preferred management one may be concerned with the specification of all three factors. It should also be noted that this rule is but one component of a management strategy which might include other feedback control rules and direct specifications of some controllable system inputs.

The specification of feedback control rules can be a difficult task in more complex situations. In some special cases optimal control methods can be used to derive feedback control rules which optimize system performance, but the presence of uncertainty greatly complicates the application of these analytical tools.² Often, then, it is necessary to specify a general form of a control rule and perform experiments to determine its parameters. A simple example related to forward contracting strategies by cash grain farmers should help to explain how a reasonable form for a feedback control rule can be determined.

Let $v(t)$ be a controllable system input which specifies the number of bushels of corn which, at time t , the operator contracts to deliver

¹In reality, of course, other variables may affect this decision.

²Optimal control techniques are discussed in detail in Aoki (1967), Karreman (1968), Sage (1968), and Kirk (1970). The first two are concerned with optimal control decisions under uncertainty.

at harvest. Let $x(t)$ be the total number of bushels contracted prior to time t --i.e. $x(t) = \sum_{i=1}^{t-1} v(i)$ --and let $d(t)$ be the total number of bushels the operator desires to have contracted at time t . The level of $v(t)$ is defined by the following expression:

$$v(t) = \begin{cases} d(t) - x(t) & \text{if } d(t) > x(t) \\ 0 & \text{otherwise} \end{cases} \quad 2.1$$

This is a feedback control rule which specifies the level of $v(t)$ at any point in time. That level is equal to the difference between desired and actual contracting levels to date. Since a contract, once made, cannot be dissolved, however, levels of $v(t)$ are restricted to non-negative values.

Actual contracting levels can be observed, but desired contracting levels cannot be.¹ In order to implement this rule, then, a more complete specification is needed, a specification which defines desired contracting levels as a function of observable variables.

The projected size of the operator's corn harvest, $h(t)$, is one factor upon which the desired level of contracting is expected to depend. Initially, then, $d(t)$ might be defined by the expression

$$d(t) = \delta h(t) \quad 2.2$$

which implies that the desired level of contracting is some specified fraction, δ , of the projected corn harvest. This is not a very satisfactory specification, however, because the operator's estimate of how many bushels of corn he expects to harvest may not change much over the period during which the rule is to be applied. Furthermore, this specification ignores the impact of prices on desired contracting levels.

¹If the operator knew his desired contracting level, he would have no need for this rule.

An additional factor which should be considered, then, is the sign and magnitude of the difference between the current forward contract price, $c(t)$, and the farmer's current estimate of the expected cash price at harvest, $e(t)$.¹ The more the contract price is above the expected cash price, the more the farmer will wish to have contracted; the more it is below the expected cash price, the less he will wish to have contracted. Therefore, a revised specification of $d(t)$ could be:

$$d(t) = h(t)[\alpha(c(t)-e(t))] \quad 2.3$$

where α is a positive constant.

The desired level of contracting may also depend on movements of the contract price. If it is rising rapidly the farmer may wish to delay the commitment of an additional portion of his crop to a forward contract. On the other hand if the contract price is falling he may wish to lock in a relatively high price. To reflect this, the specification of $d(t)$ can be further revised so that

$$d(t) = h(t)[\alpha(c(t)-e(t)) + \beta \frac{dc(t)}{dt}] \quad 2.4$$

where $dc(t)/dt$ is the rate of change in the contract price and β is a negative constant. The interaction between the two terms in brackets should be noted. If the contract price is falling but is less than the expected cash price, the first term should override the second, and no new contracting will be desired. On the other hand, if the contract price is both above the expected cash price and falling, the two terms reinforce each other to raise the desired contract level.

¹The expected cash price at harvest, $e(t)$, can be determined by an expectations model, the complexity of which can be determined by the requirements of the particular decision situation.

A third factor which could influence a forward contracting strategy is the percentage of desired corn acreage actually planted at time t , $p(t)$. Fearing unusually bad weather which could delay or prevent the planting of some of the specified acreage, some operators may hesitate to contract much of their projected harvest until planting is nearly complete. Similarly, many farm operators, fearing the consequences of a sharp downturn in prices, desire to have some of their crop contracted in nearly all situations. Therefore, the specification of $d(t)$ can be revised once again to become

$$d(t) = h(t)[\alpha(c(t)-e(t)) + \beta \frac{dc(t)}{dt} + \gamma p(t)] \quad 2.5$$

where γ is a positive constant.

One other restriction on $d(t)$ should be noted. In some situations the desired contracting level implied by the specification above may be unacceptably high either for the farm operator or for the manager of the local elevator. Therefore, it may be advisable to establish an upper bound on $d(t)$. This can be considered to be a prespecified constraint or it can be treated as a parameter. In this case the upper bound on $d(t)$ will be set at $1.5 h(t)$, which implies that a maximum of 150 percent if the projected crop can be contracted.¹

By substituting the right hand side of equation 2.5 for $d(t)$ in equation 2.1, the feedback control rule can be expressed in a form which contains only observable variables. This rule has been presented only as a relatively simple example of the types of rules that can be specified. In some situations it may be desirable to consider more

¹Due to the form of the feedback control rule no lower bound on $d(t)$ is needed. Such a lower bound could be specified, however, if necessary.

factors and to experiment further with functional form. The power of such a rule should be evident, however. Once a functional form has been specified and values for the three parameters have been selected, an adaptive marketing strategy for an entire year has been established, a strategy which only requires information which is readily available to any farm operator and which allows the decision maker to take advantage of opportunities to learn. In many decision situations such a rule may be preferable to a management strategy which specifies an inflexible marketing strategy prior to the planting season.

2.3.2 Plans, Decisions, and Actions

The preceding discussion has shown how the incorporation of feedback control rules into the specification of an action strategy introduces flexibility into the concept of what is to be decided in a particular situation. In attempting to gain a better understanding of the general characteristics of decision problems involving uncertainty and in giving further consideration to the basic question of what constitutes an action choice, it is also important to draw clear distinctions among the three related concepts of a plan, a decision, and an action. The distinctions made here parallel those made by Modigliani and Cohen (1961) and by Day (1971).

A plan is a strategy for controlling system performance through management or design which extends into an uncertain future--a strategy based on information about the current situation and on expectations concerning future events. In general plans can be altered. Such alterations may be costly, however, because resources must be expended to gather and analyze new information and to reformulate the plan

itself and because actions undertaken to implement the initial phases of a plan may restrict the opportunities open to a decision maker.¹

A decision, on the other hand, is a choice which is essentially irreversible. It is that part of a plan which is to be implemented before further planning is undertaken. As such, decisions are the desired output of the analytical process which is of primary concern in this study.

Finally, an action is the realization of a decision. In a world free of uncertainty, decisions and actions would be effectively identical. In most instances, however, events that cannot be known with certainty at the time when decisions are made affect the extent to which they can be implemented and, of course, the outcomes associated with them. What is realized may not be what was chosen. It is necessary, then, to distinguish between decisions and actions.

There is a crucial interplay among these three activities, an interplay which must be recognized during the formulation of decision problems. Though the primary focus of a decision analysis is on the choice of actions to be undertaken, one must be aware that the outcomes of current decisions affect the opportunity set which circumscribes future decisions. Therefore, it is often necessary to formulate plans for periods extending beyond the immediate period in which decisions are to be implemented. In the multiperiod model of decision making under conditions of perfect knowledge developed by Hicks in Value and Capital

¹This second point is demonstrated by Johnson (1961b) and Johnson and Quance (1972) in their discussions of asset fixity.

(1946) the planning horizon is of infinite length, since an event at any time can have an effect on the total flow of system outputs. In most cases, however, information is not perfect, and Modigliani and Cohen (1961) have observed that in an uncertain context forecasts are subject to error and reliable information about future events can be attained only at a cost--a cost which is directly related to the degree of uncertainty and inversely related to the time proximity of the future event. In such situations it is not worthwhile to formulate plans over an infinite horizon. Rather, the planning horizon should be extended only to the point where current decisions cease to be affected.¹ Furthermore, actions at the end of the planning horizon need not be planned in as great a detail as those at the beginning. In formulating decision problems, then, it is necessary to determine the appropriate length of the planning horizon.

The act of planning, which involves the collection and analysis of information, is expensive, however, and it is also important to consider how often plans should be reformulated. If the costs of planning are high or if the benefits from it are comparatively low, frequent revision of plans may not be worthwhile. Therefore, when formulating decision problems, it is also necessary to consider the length of what might be termed the decision horizon--the length of time over which decisions apply and replanning does not take place. The length of the decision horizon has an important impact on the

¹Kleindorfer and Kunreuther (1978) note that the cost of planning and the length of the planning horizon depend on the costs of forecasting future events, the degree of uncertainty, and the computational costs associated with the determination of an optimal plan.

character of decisions themselves. When it is short, the importance of learning through feedback may be minimal, and a valid decision strategy would specify levels for all controllable system inputs over the entire decision horizon. When the decision horizon is long, however, implementation of a decision strategy may involve a series of actions which are affected by factors that cannot be known with certainty at the time when decisions are made. In such instances opportunities for learning may exist, but the reformulation of plans on the basis of new information is not worthwhile. As a result, the decision may be more concerned with the choice of adaptive decision rules to be followed over the entire decision horizon than with the determination of desired levels for all choice variables over that period. Relating this to the scheme of system identification, the management process unit is viewed as a controller which directs system performance through the application of feedback control rules.

In formulating rules of this sort, one must also consider what can be termed an action horizon--the length of time between successive reassessments of the current situation and expectations for the future and reapplications of the adaptive decision rule. The action horizon corresponds to the time increment embodied in a feedback loop. Its desired length will depend on the cost of monitoring the state of the system and the environment and on the relative costs and returns of applying the feedback control rule. As such, the length of the action horizon also has an impact on the nature of decisions, and attention should be given to its specification in the process of problem formulation.

The following example should help to clarify the distinctions among plans, decisions, and actions. It should also help to demonstrate the importance of making these distinctions when formulating decision problems.

Production and investment decisions made by farmers in any given year have an impact on their operation for a number of subsequent years. They affect future cropping patterns, levels of available resources, and cash flow requirements, though their exact impact cannot be known in advance due to uncertainties with respect to a number of environmental factors beyond the control of the individual decision maker. Farmers find it desirable to formulate production and investment plans, then, but the high degree of uncertainty they face may cause them to limit their planning horizon to, perhaps, three years. Once a farmer has begun to implement his plan--once he has purchased seed, fertilizer, and other inputs needed to grow the first year's crops and, perhaps, new land or machinery--a decision has been made and in most cases extensive revision of his plan will not be worthwhile until the end of the crop year. It can be said, then, that the decision horizon is approximately one year. The farmer's decision should not be considered to be a rigid strategy which defines his actions for each day of that year, however. Rather, it is a set of specified levels for major controllable system inputs and a set of adaptive decision rules which structure future efforts to collect and analyze information and direct his actions in response to changes in the state of his operation and the environment. For example, a management strategy could be comprised in part of a set of desired acreage levels for each cropping activity and a set of adaptive rules which automatically adjust the crop plan in response to

information on current acreage planted and changes in relative prices. A simple adaptive rule might be to shift unplanted corn acreage to soybeans if the forward contract price for corn is below \$2.10 and the contract price for beans is higher than \$6.50. If such a rule were applied weekly, the action horizon would be said to be one week.

It is important to recognize the impact of each of these activities--planning, decision making, and action--on the conceptualization of the decision problem. Failure to recognize the need to plan may lead to decisions which, while beneficial in the short run, have a harmful long-run impact on the system. Similarly, it must also be recognized that planning itself is expensive and that in most cases a period of time exists over which extensive plan revision is not worthwhile. A portion of any plan, then, can be viewed as a decision which will not be altered. Finally, recognition of the fact that uncertain aspects of the environment are likely to affect the implementation of any decision and that opportunities for learning and adaptive behavior exist when the decision horizon is relatively long leads to the conclusion that in many instances decision makers should choose flexible rather than inflexible strategies.

2.4 A Formal Statement of the Decision Problem

In the analysis of complex decision problems it is often desirable to make a formal statement of the problem to be resolved. When mathematical decision aids such as those developed in this study are used in the decision analysis, this may be a necessity. In the most general terms, the basic decision problem under uncertainty can be stated in the following manner: identify a feasible action strategy which results in system performance over a specified time horizon that

can be considered optimal according to some relevant evaluative criterion. System performance is measured by system output variables, at least some of which are stochastic. An action strategy is defined by a set of controllable system input levels and/or by a set of control rules which determine the levels.¹ This is, in essence, the statement of a stochastic optimal control problem. Though it may be impossible or prohibitively expensive to find the truly optimal solution to such a problem in most decision situations, this is a generalized formal statement of the problems decision makers seek to resolve. It is a problem which requires information on the physical, human, and institutional realities of the context in which choices are made, assessments of probabilities associated with stochastic events beyond the control of the decision maker, and an understanding of the decision maker's normative beliefs if it is to be resolved successfully. As the process of problem formulation continues, a more comprehensive understanding of the problem being analyzed should be gained. This understanding of the problem should serve as a guide throughout the decision analysis.

¹Symbolically, the problem is

$$\max_{st} U = h(y(T), T) + \int_0^T g(y(t), t) dt$$

$$y(t) = f(x(t), t)$$

$$\dot{x}(t) = a(x(t), v(t), e(t), t)$$

$$r(x(t), \dot{x}(t), t) \leq \beta$$

$$s(u(t), t) \leq \gamma$$

where $y(t)$ is a vector of system output variables, $x(t)$ is a vector of system state variables, $v(t)$ is a vector of controllable system inputs, $e(t)$ is a vector of exogenous system inputs, and a is a function comprised of system design parameters. The final two constraints limit the set of allowable states and the set of allowable controllable input levels. The elements of the vector $v(t)$ are the choice variables in this problem.

2.5 An Application

In this final section, the concepts related to problem formulation presented earlier are applied to a more concrete decision problem. The example introduced here will also be used to illustrate the techniques developed in subsequent chapters. This is not an actual case example. Rather, it might be termed a synthetic case study, since it synthesizes circumstances and concerns common to many individuals' situations.

The decision maker in our example is the operator of a relatively small cash grain farm in southeastern Michigan. He owns 240 acres of tillable land on which he grows corn and soybeans. He is heavily in debt, with interest and principle repayment commitments on long and intermediate term debts of \$35,000 per year. Except for approximately \$6,000 income from off farm work by the operator and his wife, all of the family's income is derived from the farming operation. If the family's level of income is insufficient to meet debt repayment commitments and family living expenses, they face the prospect of refinancing some loans or of being forced out of farming altogether. In 1978 their income was low, and they relied in part on savings to cover expenses.

The operator views his current situation as uncertain and unsatisfactory. He feels a strong need for a higher, less uncertain level of income in the year to come. Though he has other needs and desires, this is his primary concern. The problem to be analyzed, then, is that of identifying a strategy which best provides a level of income adequate to meet debt repayment obligations and family living expenses. The operator believes this problem is a serious one, and he is willing to expend the resources required to undertake a careful analysis of his alternatives and their consequences.

our

cho

the

ces

will

of

far

off

nei

far

far

ess

is

inc

oth

dep

sys

ces

thi

as

con

con

Having identified the decision maker and the problem which motivates our analysis, it is necessary to look more carefully at the kinds of choices the farm operator can make and at the factors which will affect the outcome of these choices. One might begin by considering what processes affect the family's level of income--by defining the system that will be the focus of the analysis. In this case the system is comprised of the set of production and marketing processes which constitute the farm operation and the set of processes associated with engaging in off farm work. In order to simplify this example, it is assumed that neither the operator nor his wife is willing to take a permanent off-farm job. As a result, opportunities for affecting the pattern of off-farm earnings are limited, and the level of off-farm income will be assumed given.

The output of this system is measured by a single variable, y , which is defined as annual cash income available for family living expenses, income tax, and investment after all debt repayment commitments and other business expenses have been met.¹ The level of income realized depends on the structure of the system, on exogenous inputs to the system, and on the choices made by the operator.

The structure of the system defines more exactly the set of processes by which the system output, net cash income, is generated. In this example the conceptualization of the system will be kept as simple as possible. Standard crop budgets, which are given in Table 2.1, define

¹Other performance measures could be identified, but the cost of considering them is deemed excessive in this case, since it greatly complicates the analysis.

Table 2.1 Standard Crop Enterprise Budgets

| | Corn | | Soybeans | |
|--------------------------------|-------|-------------|----------|-------------|
| Seed, bu. | (.23) | 9.70 | (.83) | 8.30 |
| Fertilizer | | | | |
| Nitrogen, lb. | (120) | 16.80 | (10) | 1.40 |
| Phosphorous (P_2O_5), lb. | (75) | 13.50 | (50) | 9.00 |
| Potassium (K_2O), lb. | (100) | 9.00 | (25) | 2.20 |
| Lime | | .80 | | 1.10 |
| Herbicide, other chemicals | | 10.80 | | 13.00 |
| Fuel and repair | | 14.40 | | 9.60 |
| Utilities | | 2.00 | | 2.10 |
| Miscellaneous | | <u>2.20</u> | | <u>2.20</u> |
| Total Selected Cash Expenses | | 79.20 | | 48.90 |
| Drying cost, per point per bu. | | .01 | | 0 |
| Hauling cost, per bu. | | .10 | | .14 |
| Time required for | | | | |
| Planting, hours per acre | | .757 | | .757 |
| Harvest, hours per acre | | .418 | | .502 |

Source: Nott, et al. (1977).

the basic production processes for corn and soybeans. They specify not only physical inputs such as seed, fertilizer herbicides and fuel, but also the time required for planting and harvesting, the two critical fieldwork operations. These budgets alone are not considered sufficient to adequately represent crop production, however. The timeliness of planting and harvesting also affects crop yields and, ultimately, the level of income realized. Therefore, the acreage planted in each crop is classified according to when it is planted and harvested. Six planting periods and five harvest periods are defined in Table 2.2, and possible planting-harvest combinations are specified for each crop. A final characteristic of the crop production process which should be noted is the rule of thumb that all corn is planted before soybeans are planted and that all soybeans are harvested before corn is harvested. Other relevant system design features include the fact that the farmer does all the fieldwork himself, though he gets some help from his wife who hauls grain at harvest. Again to simplify the example, it is assumed that all production is sold at harvest on the cash market. Marketing alternatives such as forward contracting, hedging, storage, and participation in government set aside programs will not be considered.

These system design characteristics define the processes by which income is generated. System state variables are also useful in understanding the system's structure, since they serve as descriptors which represent fully the system's performance through time. In this example the set of state variables includes an indicator of current net cash income which is repeatedly updated as costs are incurred and crop sale receipts taken in, an indicator of ³total the acreage of each crop

Table 2.2 Planting and Harvest Periods and Possible
Crops for Each Combination

| Planting Period | Harvest Period | | | | |
|--------------------|----------------------------|---------------|---------------|---------------------------|---------------|
| | September 27- October 3 | October 4-10 | October 11-17 | October 18- November 7 | November 8-28 |
| April 25-May 10 | Corn | Corn | Corn | Corn | Corn |
| May 10-18 | Corn | Corn | Corn | Corn | Corn |
| May 19-26 | Soybeans | Corn/Soybeans | Corn/Soybeans | Corn/Soybeans | Corn |
| May 27-June 3 | Soybeans | Corn/Soybeans | Corn/Soybeans | Corn/Soybeans | Corn |
| June 4-11 | Soybeans | Corn/Soybeans | Corn/Soybeans | Corn/Soybeans | ---- |
| June 12-19 | Soybeans | Corn/Soybeans | Corn/Soybeans | Corn/Soybeans | ---- |

Source: Black, et al. (no date).

planted and harvested to date, indicators of the number of acres of each crop which remain unplanted or unharvested, and indicators of the number of bushels of each crop harvested to date. In addition, time itself is monitored, as are the number of acres of each crop planted in each of the six planting periods and the number of acres of each crop planted in a particular planting period which are harvested in each of the five harvest periods.

Inputs to the system, as well as its structure, affect the level of income realized by the farmer. A number of exogenous system inputs can be identified in this example. Those considered to be stochastic include: the price at harvest of each crop, the number of days available for fieldwork in any particular planting or harvest period, and the yield of either crop for each allowable planting-harvest combination. Subjective assessments of the probability distributions for these variables will be required for the analysis.

Relatively few controllable system inputs will be considered in this example. Since the analysis focuses on decisions related to the farming operation, those of primary concern are the number of acres rented (land rental opportunities do exist) and the number of available acres planted in each of the two crops grown, corn and soybeans. Several factors limit the range of possible values for the three controllable system inputs of primary concern. Due to the characteristics of the local land market, the number of acres rented, v_1 , can be assumed to take only five values: 0, 80, 160, 240, 320. Limits on available land imply that total crop acreage must be less than or equal to that which is owned plus that which is rented. If v_2 is acreage planted in

corn and v_3 is acreage planted in soybeans, then the following relationship must hold:

$$v_2 + v_3 \leq 240 + v_1 \quad 2.6$$

The farm operator not only has control over the several system inputs discussed above, but he also may be able to affect the performance of the system through design changes such as the purchase of new machinery or the alteration of cultural practices. Given the operator's rather precarious financial position and the fact that his crop yields have not been notably low, however, it seems best to focus the analysis on the specification of controllable system input levels. Therefore, design changes will not be considered.

What is the relevant planning horizon in this example? If herbicide carry-over problems are not considered to be of major importance, and if long term leases are not required for rental of any of the 80 acre tracts, then the choices to be made in this example have an impact on system performance and on the opportunities open to the farmer only in the year they are made. Therefore the relevant planning horizon is a single year.¹ Because the costs of the analysis to be undertaken are not insignificant, the farm operator does not wish to consider major changes in his strategy unless conditions change so dramatically that this is deemed worthwhile. The decision horizon, then, is also one year. It is recognized, however, that there will be opportunities to learn over the course of the year and that new knowledge may lead to a

¹If major design changes such as investment in new machinery were being considered, the planning horizon would need to be longer.

desire for some minor revisions in the strategy. Of particular concern are possible losses in yields due to a lack of timeliness in planting corn. Therefore, a simple feedback control rule will also be considered in the analysis. The rule takes the following form: "Regardless of specified crop acreage levels, soybeans will be planted on all unplanted acreage after v_4 (a parameter indicating a specific date)." Soybeans will not be planted before May 19, nor will corn be planted after June 3, and the operator wishes to check the feedback control rule at the end of each planting period between these dates. Given these restrictions, the possible values for the parameter, v_4 , are May 18, May 26, and June 3. The action horizon for this rule, then, is eight days during the period when it is operative.

The management strategy in this example is defined by values of the three controllable input variables and by the single feedback control rule parameter. Our problem is to find the strategy, v^* , which best satisfies the farm operator's need for a higher and more stable level of income. Constraints on allowable control variable levels and the structural characteristics of the system which determine the relationship between system inputs and outputs must be considered when the choice is made. The fact that the choice must be made under conditions of uncertainty with respect to product prices, crop yields, and time available for fieldwork must also be considered, since this means that the outcome of any strategy can be specified only in probabilistic terms. To make such a choice requires the integration of positive knowledge of what is and what may be with normative knowledge concerning the goodness or badness of particular outcomes. The expected utility

hypothesis, which states that the preferred choice of a decision maker is that which has the highest expected utility, will serve as the basis for this integration. Our problem, then, is to identify the management strategy v^* for which the associated distribution of realized net cash income, $f(y)$, maximizes the expected utility of the decision maker.

CHAPTER III

THE DETERMINATION OF SUBJECTIVE PROBABILITIES

3.1 Introduction

Choices made under uncertainty are affected by a decision maker's preferences for alternative outcomes and by his expectations concerning the likelihood of each possible outcome associated with the action strategies under consideration. Both of these factors are subjective and vary from decision maker to decision maker, and information on both is a critical input in the analysis of any decision problem in which the impact of uncertainty is of major importance. In this chapter, techniques for eliciting information on expectations and methods of structuring that information for use in a formal decision analysis will be examined.

Expectations are reflected in a decision maker's beliefs about the probabilities of different events occurring. These beliefs may be based in part on logical deductions, on inferences drawn from empirical observations, on intuition, or on a combination of all three types of information. In general, however, probabilities must be viewed as personal or subjective and, as such, cannot be judged to be correct or incorrect.¹ The problem in a decision analysis is one of representing

¹This personalistic view of probabilities whereby they are considered to be "degrees of belief" rather than objective facts has its origins in Ramsey's (1931) discussion of probability in the essay "Truth and Probability."

subjective probabilities in a manner which is consistent with the decision maker's actual beliefs and in a form which facilitates the use of this information in the evaluation of alternative choices. In many situations, this already difficult problem is made more so by the fact that the decision maker's beliefs may be poorly defined and may be based on quite limited information. As Hogarth (1975, p. 273) notes:

. . . man is a selective, stepwise information processing system with limited capacity, and, as I shall argue, he is ill-equipped for assessing subjective probability distributions. Furthermore, man frequently just ignores uncertainty.

The assessment of subjective probabilities, then, can be a difficult, complex task. Elicitation procedures should be designed to help the decision maker think in probabilistic terms. Furthermore, they should serve as an aid in structuring information from a wide range of sources, including that provided by experts more knowledgeable than the decision maker himself.

Of primary concern in the analysis of choices made under uncertainty are the subjective probability distributions of the outcomes associated with each management strategy under consideration. In general, such distributions cannot be assessed directly by the decision maker, however, since they are usually dependent both on specified levels of system control inputs and on a number of stochastic and non-stochastic environmental factors. Rather, their assessment requires both the encoding, or direct elicitation, of subjective probability distribution for important stochastic exogenous system inputs and the modelling of the relationships between system inputs--controllable and exogenous--and system outputs. The combined use of encoding and modelling allows the decision maker to break down what may be a complex

stochastic process into more manageable sub-units about which he can formulate expectations directly. It also encourages the decision maker to think explicitly about how controllable and uncontrollable factors interact to determine the outcome of any choice.

The value of careful system identification should be evident here, since it is the process by which controllable and exogenous system inputs having an important impact on system performance are identified. The number of stochastic factors considered and the complexity of the model used to represent their effort on the outcome of a particular choice should be determined by the nature of the problem being analyzed. In many instances the stochastic process under consideration will be modelled more than once with new exogenous system inputs and a more refined view of the system itself being considered at each stage. As Spetzler and Stael von Holstein (1975, p. 341-2) note:

Modelling efforts tend to be most effective and most economical if they begin with a gross model that is successively refined. A model should be refined only as long as the cost of each additional refinement provides at least comparable improvement in information. The criterion for how much information is needed depends on how significantly the information bears on the decision at hand.

The elicitation and structuring of information on subjective probability distributions, then, can be viewed as a learning process which extends the knowledge gained during system identification. It is a process which should lead to an improved understanding both of the system under consideration in a decision analysis and of the decision problem itself.

In the remainder of this chapter, procedures for encoding subjective probabilities will first be reviewed. Methods for modelling stochastic processes will then be examined with particular attention being given to the use of Monte Carlo methods and simulation techniques

to represent the performance of complex systems. Finally, these techniques are applied to the analysis of the decision problem introduced in the concluding section of the preceding chapter.

3.2 Probability Encoding Procedures

During the system identification phase of problem formulation, stochastic exogenous system inputs which are expected to have an important impact on system performance are specified. In many instances, preliminary models may be constructed and sensitivity tests performed in order to better determine which environmental factors affect system performance most significantly. Once this has been done, the encoding of subjective probabilities associated with important exogenous variables can begin.

Encoding is the process by which a decision maker's beliefs about the relative likelihood of different events are elicited and used to represent a subjective probability distribution. It is one means by which information from a range of sources is structured for use in a decision analysis. It should be emphasized once again that a single, correct subjective probability distribution does not exist. Furthermore, it should be noted that the decision maker may not even think in probabilistic terms. As Winkler (1967, p. 778) notes:

. . . there is no 'true' prior distribution. Rather, the assessor has certain prior knowledge which is not easy to express quantitatively without careful thought. An elicitation technique used by the statistician does not elicit a 'true' prior distribution, but in a sense helps to draw out an assessment of a prior distribution from the prior knowledge.

Finally, it should be noted that the encoding process may also involve the consideration of information from outside expert sources or the use

of historical data. This may be particularly important when the decision maker's own knowledge about a particular factor is limited. In such instances he may be willing to accept the assessments of others or to base his expectations for the future entirely on past occurrences.

Before actual encoding begins, the stochastic variable under consideration should be clearly defined, and its importance in the decision analysis should be recognized by the decision maker. The variable should be viewed as truly exogenous to the system under consideration. If its level will be affected by the decision maker's actions, it cannot be considered to be exogenous. Finally, the variable's relationship to other random factors should be considered carefully. If its level is conditional upon that of other exogenous variables, this should be recognized and dealt with explicitly during the encoding process.

Possible sources of bias should also be considered before encoding begins. Biases are said to exist when an encoded subjective probability distribution does not conform with the decision maker's actual beliefs. As such, they are impossible to measure, since the elicited information is the only available indicator of the decision maker's beliefs. Evidence from controlled experiments in which subjects use sample observations to assess probability distributions known to the experimenter, however, indicates that interview procedures can have an impact on responses (Hogarth, 1975). Furthermore, experiments designed to reveal how subjects assess probabilities rather than how well they assess them indicate that several of common perceptual heuristics can cause problems in the

assessment of probabilities.¹ The introduction of biases into the encoding process can be minimized through the careful design of interview procedures.

At the outset of the encoding interview, efforts should be made to gather and make note of all available information which may help the decision maker formulate his expectations. If a substantial data base of past values of the random variable being considered exists, and if the subject believes these data accurately reflect his expectations concerning future events, he may choose to let the historical data define his subjective probability distribution. This may be a reasonable procedure, for example, in the case of rainfall patterns. In other instances, this cataloging of available information may reveal that the decision maker knows little about the variable being encoded. A decision must be made, then, as to whether more information should be sought from expert sources or whether encoding should proceed. This decision will depend on the cost of new information and on the sensitivity of system performance to fluctuations in the variable under consideration.

The purpose of an encoding interview is to construct a quantitative representation of the decision maker's subjective probability distribution for a particular variable. In general, this representation takes the form of a cumulative distribution function, such as that in

¹Of these, the most notable are availability (Kahneman and Tversky, 1972), representativeness (Kahneman and Tversky, 1973), and anchoring (Tversky and Kahneman, 1974).

Figure 3.1.¹ Moments are sometimes used to describe subjective distributions, but most subjects find it difficult to translate probabilistic beliefs directly into statements of the moments of a particular distribution.

Several types of questions can be used to elicit information on a decision maker's expectations. Spetzler and Stael van Holstein (1975) classify elicitation techniques according to encoding method and response mode. They identify three encoding methods:

1. Those which require the assessor to specify a probability level while the value of the random variable is fixed--i.e., the respondent indicates the probability that the variable X will fall below x^* .
2. Those which require the assessor to specify a value of the random variable while the probability level is fixed--i.e., the respondent indicates a value of the variable X , x^* , such that the probability X will fall below x^* is equal to a specified probability level.
3. Those in which the assessor specifies both a value of the random variable and a probability value associated with it. In effect, this is done when historical data are said to reflect the decision maker's beliefs concerning future events.

Two response modes are identified: direct and indirect. Under the direct response mode the assessor is asked to explicitly specify values

¹For any value x^* of the random variable X , the corresponding value of the cumulative distribution function, $F(x^*)$, is the probability that a sample observation of X will have a value less than or equal to x^* .

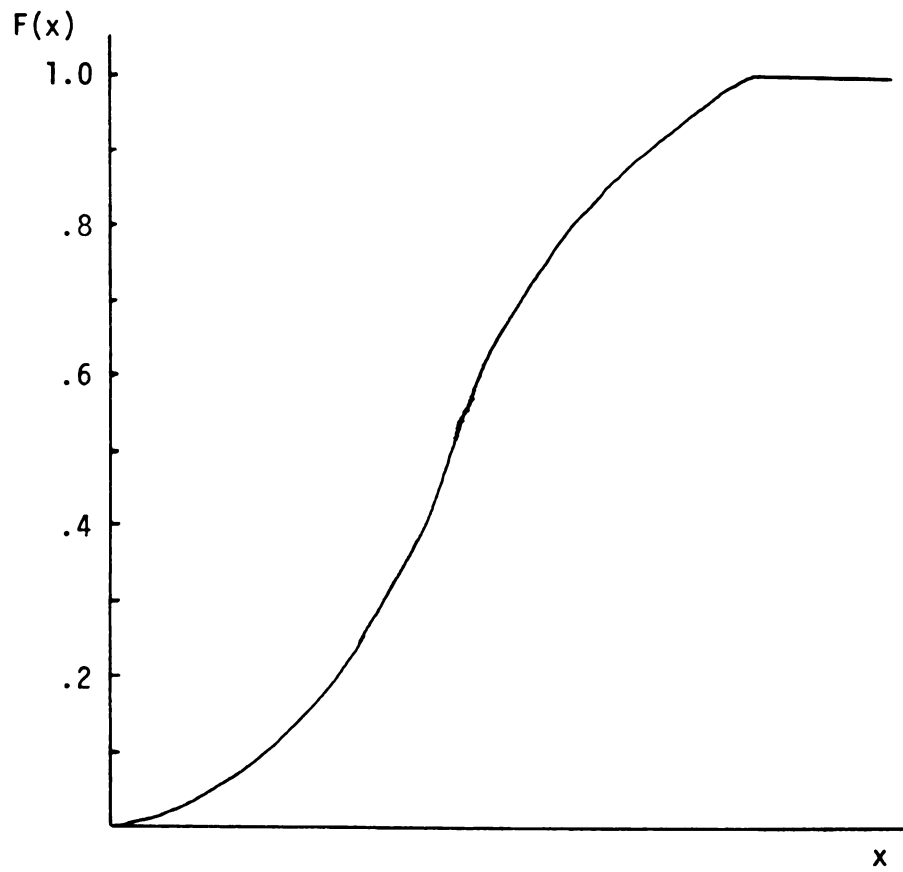


Figure 3.1 A Cumulative Distribution Function

of the random variable or probability levels which define points on the cumulative distribution function. Under the indirect mode the assessor is asked to indicate which of two or more bets he prefers. One of the bets serves as a reference which allows a probability assessment to be inferred from the assessor's response.

Several types of elicitation techniques may be used in a single encoding interview. Questioning might begin with direct response mode questioning to determine extreme values of the variable to be encoded. Indirect response mode questioning can then be used to determine a number of points on the cumulative distribution function. Finally, direct response mode questions can be asked to determine probability quartiles of the distribution which can be used as a consistency check.

At the completion of the encoding process a number of points on the cumulative distribution function of the decision maker's subjective probability distribution for the variable under consideration have been identified. Similarly, if historical data are used in lieu of subjective assessments to define the distribution, each observation can be assigned a position on a cumulative distribution function according to the following rule: "If a sample of n observations is drawn from some distribution and arranged in order of size, the k th observation is a reasonable estimate of the $k / (n + 1)$ fractile of the distribution" (Schlaiffer, 1959, p. 104). The cumulative distribution function must be defined for all possible values of the random variable, however, not simply at selected points. This is usually accomplished by sketching a smooth curve through the observed points, though if there are good reasons to believe the distribution of a random variable is from a

particular family of distributions with cumulatives of a known form, regression techniques can be used. In either case the reliability of such a representation may be questioned, especially when only a few data points have been identified.

Anderson (1974b) has investigated the estimational reliability of hand-sketched smoothed cumulative distribution functions based on sparse empirical data. The questions he poses are also relevant in connection with the construction of smoothed cumulative distribution functions from a small number of subjectively assessed points. His results indicate that, as expected, estimational reliability increases when a large number of observations are available, but they also show that in a surprisingly large number of cases a fairly good estimate of an underlying distribution can be made on the basis of only three to five observations. These results are somewhat encouraging, but they should also serve as a warning of the need in some cases to consider explicitly the inexactness of assessed probability distributions, as is done by Watson, Weiss and Donnell (1979).

The encoding interview should end with the decision maker's verification of the quantitative representation of his beliefs. This can be done by asking him to examine either the cumulative distribution function which has been constructed or a random sample drawn from the distribution it defines. More formal verification procedures involve the use of scoring rules (Winkler, 1969; Stael von Holstein, 1970; Savage, 1971),

but as Hogarth (1975) notes, the usefulness of such rules is questionable in many situations.¹

Finally, it should be noted that the encoding procedures discussed above are designed for the determination of subjective probability distributions of random variables judged to be statistically independent of all other random factors under consideration. In many instances, however, random factors are not independent. On a particular farm, for example, yields for two crops such as corn and soybeans are likely to be highly correlated. Similarly, prices received for these two crops would not be expected to be statistically independent. In such a situation one of two alternative procedures can be followed: the process by which the correlated random variables are determined can be modelled back to the point where all stochastic exogenous factors can be assumed to be independent or the decision maker's joint probability distribution for the correlated variables can be assessed directly. Neither alternative is particularly attractive. Modelling can be costly, and as the model becomes more complex the number of variables to be encoded may increase rapidly. Furthermore, while the decision maker may have well formulated expectations about many of the non-independent random variables, he may have little or no knowledge of the statistically independent underlying variables in the more extensive

¹A scoring rule is a payoff function with the vector of stated probabilities for each of a set of mutually exclusive and expansive events and a vector of probabilities representing the decision maker's true beliefs being the arguments. If a scoring rule is strictly proper, it will be maximized when the stated probabilities coincide directly with the assessor's true beliefs.

model. Encoding of joint probability functions, on the other hand, is a difficult, time-consuming process. Experimental results indicate that many subjects are ill-equipped for the assessment of correlations between random variables (Chapman, 1967; Tversky and Kahneman, 1973). Therefore, elicited distributions may simply reflect poorly formulated expectations.

When joint specification rather than more extensive modelling is the preferred course of action, it is advisable to rely on historical data whenever possible. For example, if distributions for rainfall and daylight hours without cloud cover over a particular two-week period are to be assessed, past weather data could probably be used to represent most decision makers' expectations. Similarly, yield data for several crops over an extended time period, if corrected for time trends and other identifiable factors, may provide adequate information to construct a marginal distribution for each crop and to estimate correlation coefficients between crops. Even in the case of crop prices for which past experience may not be relevant in the formulation of each marginal distribution, it may be possible to use historical data to estimate correlation coefficients which could be used in conjunction with marginal distributions determined by other methods.

In cases where historical data are not available or are considered to be irrelevant, joint specification of bivariate subjective probability distributions can be accomplished by encoding one of the marginal distributions and then encoding conditional distributions for the second variable at several values of the first. Anderson, Dillon, and Hardaker (1977) describe this procedure in some detail and explain how it can be extended to cases where more than two correlated variables are to be encoded.

3.3 Modelling Stochastic Processes

The encoding techniques reviewed in the preceding section are used to elicit direct assessments of subjective probability distributions. In most decision situations, however, such direct probability assessments can be made only for the distributions of exogenous system input variables. They usually cannot be made for the distributions of system output variables, the distributions of primary concern in the evaluation of alternative choices. Rather, these distributions, which depend on complex interactions among a number of factors, must be determined indirectly by modelling system performance. A model is a deterministic mathematical representation of the set of processes by which controllable and exogenous system inputs determine system output levels. As will be demonstrated below, given information about the levels of controllable system inputs which define a particular strategy and information about the probability distributions of exogenous system inputs, a model can be used to determine the associated probability distribution of system output levels.

Even in situations where the distributions of system output variables associated with alternative strategies can be assessed directly, a model of the system under consideration can be useful for several reasons. First, it can increase the decision maker's understanding of the set of processes which determine the outcome of any strategy, since modelling can be viewed as a learning activity. Second, by providing a logical representation of the processes which comprise a system, a model allows the decision maker to focus his attention on the formulation and representation of expectations about future levels of

individual exogenous system input variables. He need not consider all such factors and their interactions with other determinants of system performance simultaneously. Finally, if the number of alternatives being considered is large relative to the number of exogenous system input variables, the use of a model to determine system output variable distributions can significantly reduce the number of probability distributions which must be encoded, since only the distributions of exogenous system input variables must be assessed.¹

The exact nature and complexity of the model used in any particular decision analysis will depend on the characteristics of the problem under consideration. In some instances the appropriate model may be quite simple. For example, if the set of controllable inputs is defined by v , a column vector of acreage levels for each of several crops; if the set of exogenous system inputs is defined by e , a column vector of net returns per acre for each crop activity; and if the total net return for all crop activities, y , is the only system output variable of concern, the appropriate model of this system may simply be the following linear equation:

$$y = e'v \quad 3.1$$

In other cases, much more complex models may be required to adequately represent the relationships among system inputs and system outputs. This is true, especially, when the system under consideration is dynamic and when strategies are defined by feedback control rules as

¹The same probability assessments for these variables are used in the determination of system output distributions for each strategy considered. Therefore, exogenous system inputs must be truly exogenous--i.e., their levels must not be significantly affected by system performance.

well as by fixed specifications of controllable system input levels.¹ Specific conceptual and quantitative tools used in systems modelling will not be discussed in this study. Excellent discussions of such techniques can be found in Forrester (1961), Manetsch and Park (1977a), and Manetsch (1978).

As defined above, a model is a deterministic representation of the relationship between a set of system inputs and a set of system outputs. Given specified levels of all system inputs, controllable and exogenous, the set of system outputs can be calculated exactly. In decision situations involving uncertainty, however, levels of stochastic exogenous system inputs cannot be known exactly prior to their occurrence. This implies, of course, that the system output levels associated with a particular strategy cannot be determined exactly either. Rather, they can be specified only in probabilistic terms. In such instances a system model, despite its deterministic character, can be of use in describing system performance.

In some special cases, a system model can be the basis for the analytical determination of the distribution of system outputs associated with any strategy being considered. Returning to the simple linear model specified in equation 3.1, for example, if each random factor in the vector of exogenous system inputs, e , is normally distributed, the distribution of total net revenue, y , is also normal with mean, μ , and variance, σ^2 , defined by the following expressions:

¹See Forrester (1961), Cyert and March (1963), and Dent and Anderson (1971) for examples of more complex models.

$$\mu = m'v$$

$$\sigma^2 = v'\Omega v$$

3.2

where m is a column vector of the expected net revenues for each crop activity and Ω is the variance-covariance matrix for net returns.

Anderson (1975) has shown that the distribution of y can also be determined analytically when each element of e has a Beta distribution.

In situations where subjective probability distributions for exogenous system inputs are not all members of the same family of distributions or where a more complex model is required to represent system performance, it may not be possible to analytically derive distributions for system outputs. In fact, when a model is particularly complex, it may not even be possible to calculate system outputs analytically for the special case when levels for all controllable and exogenous system inputs are known with certainty. In such instances, numerical simulation techniques and Monte Carlo methods are required to determine system output distributions.

Manetsch and Park (1977b, p. 8-1) define simulation as "a technique for obtaining particular time solutions of a mathematical model corresponding to specific assumptions regarding model inputs and values assigned to parameters." The model specified in equation 3.1 can provide the basis for a simple example of simulation. Consider the case in which 200 acres of land are to be planted and only two crop activities are possible. Let corn be crop 1 and soybeans be crop 2. If the controllable and exogenous system input vectors, v and e respectively, are defined as follows,

$$v = \begin{bmatrix} 150 \\ 50 \end{bmatrix}$$

$$e = \begin{bmatrix} 95.00 \\ 100.00 \end{bmatrix}$$

3.3

then simulation of system performance for this particular case requires only that the following matrix multiplication be carried out:

$$y = [95.00 \quad 100.00] \begin{bmatrix} 150 \\ 50 \end{bmatrix} \quad 3.4$$

In this case, with 150 acres of corn and 50 acres of soybeans planted, total net revenue, y , is equal to \$19,250.

Systems of concern in most practical decision situations are larger and more complex than this one, and their simulation is generally more involved. Frequently numerical solution techniques are required, and many simulation models are computerized. One of the distinct advantages of simulation, however, is that it is a remarkably flexible procedure which allows complex processes to be represented realistically. Naylor, et al. (1966), Schmidt and Taylor (1970), and Manetsch and Park (1977b) all provide excellent discussions of simulation techniques.

Monte Carlo methods are commonly used in combination with simulation to model the performance of complex stochastic systems. Under this approach, numerical procedures are employed to generate sample observations from the decision maker's subjective probability distributions for exogenous system input variables. Each sample vector, e^* , specifies a level for each exogenous system input and, as such, defines a state of the system's environment. By constructing a large number of sample states of the environment and simulating the system performance associated with a particular strategy for each of these environmental states, a set of sample observations from the distribution of system outputs associated with that strategy is generated. These observations can be used to

define a cumulative distribution function according to the procedures described in the discussion of encoding.¹

Returning once again to the simple example discussed above, let the joint subjective probability distribution of the two elements of the exogenous systems input vector, e , be defined in the following manner. The marginal distribution of revenues for corn is normal with mean \$115.00 and standard deviation \$35.00. That for soybeans is a member of the gamma family of distributions with a mean of \$135.00, a standard deviation of \$40.00 and a minimum value of \$55.00. The correlation coefficient for the two net revenues is .75. Using Monte Carlo techniques described in detail in Appendix A, the following five sample vectors from this joint probability distribution were generated:

$$\begin{aligned}
 e^1 &= \begin{bmatrix} 107.87 \\ 107.43 \end{bmatrix} & e^2 &= \begin{bmatrix} 156.77 \\ 128.99 \end{bmatrix} & e^3 &= \begin{bmatrix} 80.08 \\ 92.23 \end{bmatrix} \\
 e^4 &= \begin{bmatrix} 158.18 \\ 147.11 \end{bmatrix} & e^5 &= \begin{bmatrix} 152.06 \\ 210.51 \end{bmatrix} & & 3.5
 \end{aligned}$$

They represent five sample states of the environment. Simulating system performance under the strategy defined by the controllable system input vector $v' = [150 \ 50]$, five sample observations from the distribution of net income levels associated with this strategy are determined:

$y^1 = \$21552.00$, $y^2 = \$29965.00$, $y^3 = \$16623.50$, $y^4 = \$31082.50$,
 $y^5 = \$33334.50$. These five sample observations were used to construct the cumulative distribution function shown in Figure 3.2 In general,

¹In addition to Schlaiffer (1959), Mood and Graybill (1963) and Barnett (1975) also discuss the validity of the rule which, when N observations of a random variable are arrayed in increasing order, states that the K th observation can be used as an estimate of the $K/(N+1)$ fractile.

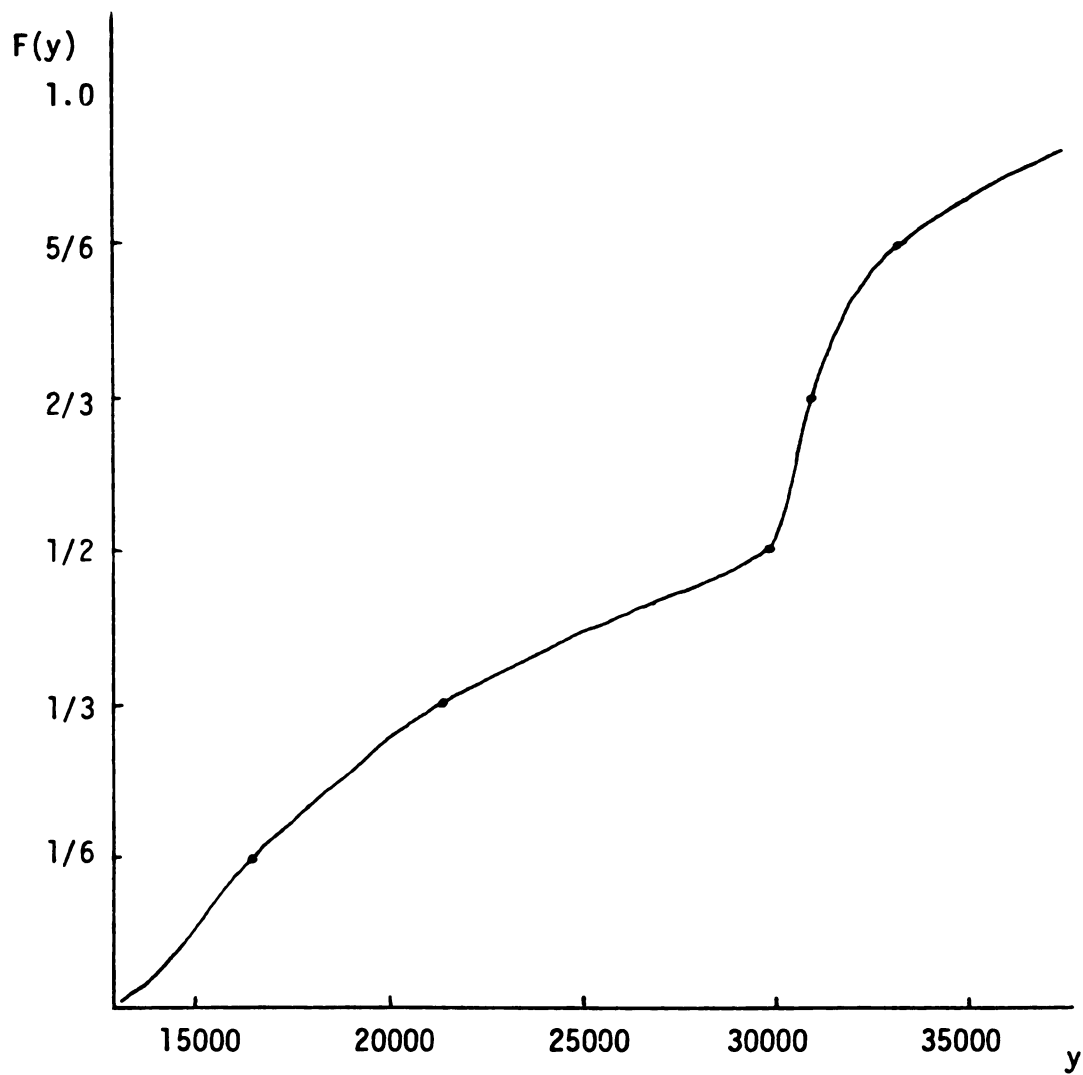


Figure 3.2 A Cumulative Distribution Function
Based on Five Sample Observations

at least 20 sample environmental states should be simulated, and for many problems it may be desirable to simulate as many as 50 to 100 sample states.

Even this simple example demonstrates the power of this approach. Because the marginal distributions of the two net return variables are not of the same family, the distribution of net income levels cannot be derived analytically. The combinations of Monte Carlo sampling techniques and simulation, however, permits the representation of the distribution. This same approach can be easily extended for use in the analysis of much more complex systems in which the interactions between controllable and exogenous system inputs are not so straightforward.¹ It should also be noted that this method imposes no restrictions on the nature of the system inputs or system outputs. Subjective probability distributions for environmental factors can take a form which most closely reflects the decision maker's encoded beliefs. System output distributions are determined by the structure of the model, by the subjective probability distributions for exogenous system inputs, and by the management control strategy. When used in a decision analysis, these distributions can be described by their moments or by their cumulative distribution functions.

One serious criticism of this approach to modelling stochastic processes is that statistical dependence between random environmental factors is often ignored (Anderson, 1974a). This may be due, in part, to difficulties associated with the joint specification of subjective

¹The example in the final section of this chapter demonstrates how this approach can be applied in the analysis of a more complex system.

probability distributions. Even when statistical dependence has been recognized in the encoding process, however, it is often ignored in stochastic system models due to a lack of available techniques for generating sample observations from multivariate probability distributions.

Procedures have been developed for the generation of random variates from a wide range of univariate probability distributions (Naylor, et al., 1966; Schmidt and Taylor, 1970). Process generators have also been formulated for several multivariate distributions, most notably the multivariate normal and Wishart distributions (Naylor, et al., 1966; Newman and Odell, 1971). More recently, Coleman and Saipé (1977) have developed a procedure for generating serially correlated lognormal variates, which is a special case of more general procedures for modelling bivariate random variables with prescribed marginals and correlations (Coleman and Saipé, 1976). A need remains, however, for a generalized multivariate process generator which permits greater flexibility concerning the specification of marginal distributions and which has the capacity to be easily extended beyond the bivariate case. Such a procedure has been developed as part of this study and is described in detail in Appendix A.¹ This generalized multivariate process generator can be used to generate sample observations from multivariate distributions comprised of up to fifty random variables.² The

¹Procedures used to generate sample observations from univariate distributions are also reviewed.

²The program can easily be expanded to model processes with still more individual variates.

marginals of the distribution modelled can be of any form and they need not all belong to the same family of distributions. It is assumed only that enough information is available on each marginal to construct its cumulative distribution function and that the correlation coefficient between each pair of variates within the distribution can be specified. The only restriction placed on the matrix of correlation coefficients is that it be positive-definite and symmetrical, a condition required for feasibility and internal consistency.

The existence of such a procedure greatly enhances the power of the approach to the modelling of stochastic processes described in this section. It permits greater realism in the representation of underlying probability distribution without requiring that the stochastic dependence between some random factors, which has an important impact on choices in many decision situations, be ignored.

3.4 An Application

The techniques introduced in this chapter can be applied to the cash grain farm example formulated in Chapter II. Crop yields for each planting-harvest period combination, product prices, and time available for fieldwork have been identified as stochastic factors which have an important impact on the level of income realized by the farm operator. In this section the specification of subjective probability distributions for each of these exogenous system inputs and the use of simulation to determine the impact of these random factors on the outcomes associated with particular management strategies are discussed.

Because the case farm used in this example is a synthetic one, no actual decision maker has been identified. Therefore, experts in the Department of Agricultural Economics at Michigan State University were relied on for the assessment of probability distribution for the stochastic factors. This is not altogether unrealistic, since in many cases actual decision makers choose to rely heavily on the opinions of experts in the formulation of their expectations.

The assessment of probability distributions for yields was based in part on historical data and in part on more subjective information. Estimates of expected corn and soybean yields for each planting-harvest combination are given in Tables 3.1 and 3.2.¹ These estimates are based on figures used in Telplan Program #18, a commonly used decision aid which focuses on choices similar to those being analyzed in this example. They are the product of a group assessment of experimental data and the personal observations of experts. No estimates of variances or other features of these probability distributions were made by this group; nor did they assess the degree of correlation between yields for different crops and different planting-harvest combinations. Therefore, the following subjective assessments were made. All yield distributions were assumed to be normal, having means equal to those specified in Tables 3.1 and 3.2. Specification of the variances of these distributions was based on the assumption that the coefficient of variation for all corn yields is 11 percent and that for all soybean yields is

¹Base yields of 100 bu/acre for corn and 33 bu/acre for soybeans are assumed.

Table 3.1 Average Corn Yield and Moisture Content
by Planting and Harvest Date

| Planting Period | Harvest Period | | | | |
|--------------------|----------------------------|--------------|---------------|---------------------------|---------------|
| | September 27- October 3 | October 4-10 | October 11-17 | October 18- November 7 | November 8-28 |
| April 25-May 10 | 90 bu @ 28% | 100 bu @ 28% | 99 bu @ 26% | 98 bu @ 23% | 88 bu @ 21% |
| May 11-18 | 82 bu @ 30% | 92 bu @ 30% | 91 bu @ 28% | 90 bu @ 25% | 80 bu @ 23% |
| May 19-26 | None | 84 bu @ 32% | 83 bu @ 30% | 82 bu @ 27% | 71 bu @ 25% |
| May 27-June 3 | None | 76 bu @ 35% | 75 bu @ 33% | 74 bu @ 30% | 62 bu @ 27% |
| June 4-11 | None | None | None | None | None |
| June 12-19 | None | None | None | None | None |

Source: Black, et al. (no date).

Table 3.2 Average Soybean Yield by Planting and Harvest Date

| Planting Period | Harvest Period | | | | |
|-----------------|------------------------|--------------|---------------|-----------------------|---------------|
| | September 27-October 3 | October 4-10 | October 11-17 | October 18-November 7 | November 8-28 |
| April 25-May 10 | None | None | None | None | None |
| May 11-18 | None | None | None | None | None |
| May 19-26 | 33 bu | 31 bu | 29 bu | 22 bu | None |
| May 27-June 3 | 32 bu | 30 bu | 28 bu | 21 bu | None |
| June 4-11 | 31 bu | 29 bu | 26 bu | 20 bu | None |
| June 12-19 | 27 bu | 25 bu | 23 bu | 16 bu | None |

Source: Black, et al. (no date).

7.5 percent.¹ Concerning yield correlations, the correlation coefficient between any two corn yield distributions or any two soybean distributions was set equal to .90, while that between any pair of distributions comprised of one corn yield distribution and one soybean yield distribution was set at .80.² These parameters--means, coefficients of variation, and correlation coefficients--define a multivariate normal distribution. Since there are thirty-four individual yield distributions, eighteen for corn and sixteen for soybeans, this multivariate distribution is comprised of thirty-four random variables. One additional characteristic was specified for this set of distributions. It was felt that the multivariate normal distribution, as specified, did not adequately account for the possibility of extremely low yields due to serious drought. Such conditions occur in southeastern Michigan about one year in twenty. Therefore it was specified that in any year there is a .05 probability that drought conditions will prevail and that corn yields will be one-half and soybeans two-thirds of what they would have been under more normal conditions.

¹The coefficient of variation, C, is defined by the expression

$$C = \frac{\sigma}{\mu}$$

where σ is the standard deviation of a distribution and μ is its mean. If the mean and coefficient of variation are known, the variance of that distribution, σ^2 , is defined by the expression $\sigma^2 = (C\mu)^2$.

²The correlation coefficient for any two random variables x only is defined by the expression

$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

where σ_{xy} is the covariance between x and y and σ_x and σ_y are the respective standard deviations of x and y.

Assessments of probability distributions of time available for fieldwork in each planting or harvest period were based on the information presented in Table 3.3, which is also based on figures used in Telplan Program #18. All distributions of time available for fieldwork were assumed to be members of the Beta family of distributions. This assumption was made because of the flexibility of Beta distributions and because, like the amount of time available for fieldwork, Beta distributions are bounded from above and below. The choice of parameters for each distribution was based solely on the information in Table 3.3, which is adequate to determine upper and lower bounds and one intermediate point on the cumulative distribution function. With the aid of tables of twentiles of the standard Beta distribution given in Pratt, Raiffa, and Schlaiffer (1965), parameters were selected for each period according to the simple criterion that the cumulative distribution function should pass as close as possible to the single observed data point. These parameters are given in Table 3.4. No information on correlations between time available for fieldwork in different periods was available, but it was felt that correlations do exist between levels observed in adjacent or nearly adjacent time periods. Therefore the following assumptions were made. The correlation coefficient for the time available in any two adjacent periods was set at .5; that for periods separated by a single period was set at .3; and that for periods separated by two periods was set at .1. All other correlation coefficients were set at 0, including those between any planting period and harvest period. => probably any

Encoding procedures such as those outlined in Section 3.2 above were used to elicit the author's own subjective probability distributions

**Table 3.3 Percent of Time Available for Fieldwork by
Calendar Period for Well Drained Sandy
Loom Soils in the Lenawee, Monroe,
Livingston County Area**

| Period | Calendar Days | Percentage ^a |
|------------------------|------------------|-------------------------|
| April 25-May 10 | 16 | 50 |
| May 11-18 | 8 | 37 |
| May 19-26 | 8 | 65 |
| May 27-June 3 | 8 | 70 |
| June 4-11 | 8 | 70 |
| June 12-19 | 8 | 70 |
| September 27-October 3 | 7 | 53 |
| October 4-10 | 7 | 53 |
| October 11-17 | 7 | 53 |
| October 18-November 7 | 21 | 33 |
| November 8-28 | 21 | 14 |

^aThe probability that the percentage of days available for fieldwork will be less than this value is .3.

Source: Black, et al. (no date).

Table 3.4 Beta Distribution Parameters^a for Time Available
for Fieldwork by Calendar Period

| Period | Lower Bound | Upper Bound ^b | α | β | Mean ^c | Standard Deviation ^d |
|-----------------|----------------|-----------------------------|----------|---------|-------------------|------------------------------------|
| April 23-May 10 | 0 | 192 | 9 | 7 | 108 | 21.1 |
| May 11-18 | 0 | 96 | 7 | 9 | 42 | 11.5 |
| May 19-26 | 0 | 96 | 14 | 6 | 67 | 9.6 |
| May 27-June 3 | 0 | 96 | 12 | 4 | 72 | 10.1 |
| June 4-11 | 0 | 96 | 12 | 4 | 72 | 10.1 |
| June 12-19 | 0 | 96 | 12 | 4 | 72 | 10.1 |
| Sept. 27-Oct. 3 | 0 | 84 | 10 | 7 | 49 | 9.7 |
| October 4-10 | 0 | 84 | 10 | 7 | 49 | 9.7 |
| October 11-17 | 0 | 84 | 10 | 7 | 49 | 9.7 |
| Oct. 17-Nov. 7 | 0 | 252 | 6 | 9 | 101 | 30.9 |
| November 8-28 | 0 | 252 | 3 | 12 | 50 | 25.2 |

^aThe density function of the standard Beta distribution is given by the expression:

$$f(x) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1} & \text{for } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

^bTime is measured in hours under the assumption of twelve-hour work days.

^cThe mean of the Beta distribution is given by the expression:

$$\mu = \frac{\alpha}{(\alpha+\beta)} (b_2 - b_1) + b_1$$

where b_2 and b_1 are the upper and lower bound values.

^dThe standard deviation of the standard Beta distribution is given by the expression:

$$\sigma = \sqrt{\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} (b_2 - b_1)^2}$$

where b_2 and b_1 are again the upper and lower bound values.

for corn and soybean prices at harvest in 1979. These assessments, which are representative of expectations in early January 1979, were reviewed by outlook specialists for these crops and revised to incorporate their opinions. Upper and lower bounds beyond which each price level could not reasonably be expected to fall and three intermediate points as each distribution were specified. As was the case with days available for fieldwork distribution, it was deemed reasonable to represent these probability assessments with Beta distributions, and Beta tables were again used to select appropriate parameters. The price distribution parameters are given in Table 3.5. Historical data series for November prices of corn and soybeans over the period 1958-1972 were used to assess the degree of correlation between the two prices.¹ The sample correlation coefficient on de-trended prices for this period is .56. Because some experts feel there is less basis for correlation between corn and soybean prices now than in the past, the subjective estimate was set at .45. Finally, production of corn and soybeans in the region surrounding southeastern Michigan does not represent a major portion of national production of either crop. Therefore, both prices were assumed to be statistically independent of yields and days available for fieldwork.

Clearly the determination of subjective probability assessments can, in practice, be an imprecise undertaking, though the resulting distributions in this instance are not unreasonable. This high degree of imprecision leads us to two important observations concerning the

¹The period after 1972 was considered to be somewhat atypical of the current situation due to the lack of government stocks from 1973-76.

Table 3.5 Beta Distribution Parameters for Corn
and Soybean Price Distributions^a

| Harvest Price of Corn | Lower Bound | Upper Bound | α | β | Mean | Standard Deviation |
|--------------------------|-------------|-------------|----------|---------|--------|-----------------------|
| Corn | \$1.80 | \$3.50 | 2 | 8 | \$2.14 | \$.21 |
| Soybeans | \$4.50 | \$8.00 | 2 | 2 | \$6.25 | \$.78 |

^aSee Table 3.4 for explanation of parameters.

analysis of decisions made under uncertainty. First, we note that there is a lack of available material which can serve as an aid in the assessment of probability distribution. The need for such information is generally not considered when agronomic experiments are designed or when outlook information is reported. When one considers the degree of uncertainty experienced by agricultural producers, however, and the impact of this uncertainty in the decision process, the need for more expert assessments of probability distribution is evident. Second, it should be noted that decisions must be made even in the absence of reliable information upon which to base probability assessments. The costs of obtaining additional information must be weighed against the possible benefits. Further refinements should not be made in subjective probability assessments beyond the point at which the decision maker and the analyst believe the farmer's expectations are reasonably well represented.

To this point probability distributions for exogenous system inputs judged to have a significant impact on system performance, as measured by net cash income, have been specified. The task of actually determining the effect these factors have on the distribution of net cash income levels remains. This requires the application of the Monte Carlo simulation techniques outlined above in Section 3.3. For any management strategy under consideration, net cash income levels are determined for a number of randomly selected states of nature. These levels are viewed as sample observations from the distribution of net cash income levels under the particular management strategy. When arrayed in order of increasing magnitude they serve as the basis for the construction of the cumulative distribution function of that distribution.

In this example, it will be recalled, a management strategy is defined by levels of control variables indicating acreage rented, acreage to be planted in corn, and acreage planted to be planted in soybeans and by a single feedback control role parameter indicating the date after which all unplanted acreage is to be planted in soybeans. A state of the environment is defined by a set of specified values for all exogenous system inputs--by values for all non-stochastic environmental factors and by one sample observation from the multivariate probability distribution comprised of crop yields, time available for fieldwork, and product prices.

A simple simulation model was specified to determine the level of net income realized under any particular managerial strategy in a given state of the environment. The simulation begins with the computation of charges for land rental, if any. Subject to time available for fieldwork, the model then simulates the planting of corn until the specified corn acreage is attained or until the date after which all remaining acreage is to be planted in soybeans. Planting of soybeans then proceeds until all acreage is planted or until June 19, the final day of the last planting period. Throughout the planting process, system state variables indicating the acreage of each crop planted in each planting period are repeatedly updated. There is no assurance that all available acres will be planted in a particular state of nature; this depends on levels of time available for fieldwork. Costs for seed, fertilizer, herbicides, and fuel are incurred for each acre actually planted.

Harvesting is simulated in a similar manner. Subject to time available, soybeans are harvested as quickly as possible, with acreage

planted first being harvested first. This continues until all planted soybeans acreage is harvested or until November 8, the date after which unharvested acreage is judged to be a total loss. The harvest of corn then begins, again with acreage planted first being harvested first. This process continues until all planted corn acreage is harvested or until November 28, the last day of the final harvest period. Again, there is no assurance that all acres planted will be harvested.¹ All harvested acreage is classified according to crop, planting period, and harvest period, and system state variables indicating the number of acres in each category are repeatedly updated. The values of these variables are multiplied by corresponding crop yields for each planting-harvest combination to determine the total number of bushels of each crop harvested. Drying and hauling costs are assessed for each bushel harvested. Finally, receipts from crop sales are determined by multiplying the number of bushels of each crop harvested by the relevant price, and net cash income is computed by subtracting costs incurred and debt repayment commitments from the sum of crop receipts and off-farm income.

This model was used to determine net cash income levels realized in twenty randomly selected states of nature for each of the two management strategies defined in Table 3.6. Levels for stochastic factors in each state of nature were generated using the Monte Carlo procedures described in Appendix A. In effect each state can be viewed as a sample observation from the combined multivariate probability distribution of prices, yields, and days available for fieldwork. Net income levels realized under Strategy 1 are given in the first column of Table 3.7.

¹ Nothing is recovered from unharvested soybean acreage, but yields equal to one-half those realized for corn planted in the fourth planting period and harvested in the final harvest period are assumed to be recoverable on unharvested corn acreage.

Table 3.6 Two Possible Management Strategies

| Central Variable | Strategy | |
|---------------------------------|----------|--------|
| | 1 | 2 |
| Land Rental | 0 | 240 |
| Corn Acreage | 0 | 180 |
| Soybean Acreage | 240 | 300 |
| Stopping Date for Corn Planting | June 3 | May 26 |

Table 3.7 System Performance under Strategy 1

| Cash Income | Corn | | | Soybeans | | | | |
|----------------|------------------|--------------------|------------------|----------|------------------|--------------------|------------------|-------|
| | Acres Planted | Acres Harvested | Average Yield | Price | Acres Planted | Acres Harvested | Average Yield | Price |
| 6551.16 | 0.00 | 0.00 | 1 ^a | 2.07 | 240.00 | 240.00 | 32.90 | 6.17 |
| 3614.65 | 0.00 | 0.00 | 1 | 2.19 | 240.00 | 240.00 | 32.91 | 5.67 |
| -2754.79 | 0.00 | 0.00 | 1 | 2.14 | 240.00 | 240.00 | 23.29 | 6.20 |
| 8054.64 | 0.00 | 0.00 | 1 | 2.07 | 240.00 | 240.00 | 31.07 | 6.68 |
| 4134.00 | 0.00 | 0.00 | 1 | 2.30 | 240.00 | 240.00 | 30.86 | 6.15 |
| 4601.17 | 0.00 | 0.00 | 1 | 2.13 | 240.00 | 240.00 | 30.19 | 6.30 |
| 2540.96 | 0.00 | 0.00 | 1 | 2.07 | 240.00 | 240.00 | 29.74 | 6.20 |
| 3029.85 | 0.00 | 0.00 | 1 | 2.22 | 240.00 | 240.00 | 29.75 | 6.26 |
| -11875.09 | 0.00 | 0.00 | 1 | 1.87 | 240.00 | 240.00 | 27.20 | 4.55 |
| 20347.68 | 0.00 | 0.00 | 1 | 2.03 | 240.00 | 240.00 | 32.73 | 7.91 |
| 13452.98 | 0.00 | 0.00 | 1 | 2.31 | 240.00 | 240.00 | 30.24 | 7.63 |
| 15182.27 | 0.00 | 0.00 | 1 | 2.31 | 240.00 | 240.00 | 32.27 | 7.35 |
| -7825.05 | 0.00 | 0.00 | 1 | 1.83 | 240.00 | 240.00 | 27.64 | 5.05 |
| -7637.09 | 0.00 | 0.00 | 1 | 2.39 | 240.00 | 240.00 | 20.42 | 6.88 |
| 7129.61 | 0.00 | 0.00 | 1 | 2.01 | 240.00 | 240.00 | 26.10 | 7.77 |
| -1084.70 | 0.00 | 0.00 | 1 | 2.20 | 240.00 | 240.00 | 27.81 | 6.07 |
| -2547.49 | 0.00 | 0.00 | 1 | 1.91 | 240.00 | 240.00 | 29.02 | 5.41 |
| 349.47 | 0.00 | 0.00 | 1 | 2.22 | 240.00 | 240.00 | 31.66 | 5.54 |
| 43072.31 | 0.00 | 0.00 | 1 | 2.20 | 240.00 | 240.00 | 32.22 | 7.09 |
| 19111.57 | 0.00 | 0.00 | 1 | 2.32 | 240.00 | 240.00 | 31.25 | 7.15 |

^aThe average corn yield is undefined when no corn is planted. $\mu = \$3815.71$ $\sigma = 8357.33$

Information on soybean prices and average yields per acre is also given. If this strategy is selected by the operator of the farm in our example the expected level of net cash income is rather low--only \$3,816. The probability is .3 that net cash income will be negative, but there is only a probability of .05 that losses will exceed \$10,000. On the other hand, there is a probability of .25 that net cash income will exceed \$10,000 and a probability of .05 that it will exceed \$20,000. Net income levels realized under Strategy 2 are given in Table 3.8. Under this strategy, which calls for the rental of 240 acres and a more balanced crop mix, the expected level of net cash income is much higher--\$10,798. The probability is .6 that net cash income will exceed \$10,000 under this strategy and .3 that it will exceed \$20,000. The probability of realizing a negative net income level is .25, which is less than under Strategy 1. When losses occur, however, they tend to be substantial and there is a .15 probability that net cash income will be less than \$10,000. The other information in Table 3.8 demonstrates how the management strategy is revised by the feedback control rule. In ten of the twenty states of nature ¹⁸⁰ less than 190 acres of corn are planted because of the stipulation that all unplanted acreage be planted in soybeans after May 26. The information in this table also demonstrates that when the number of acres cultivated reaches this high a level, there is no assurance that all available acreage will be planted or that all planted acres will be harvested.

The figures given in Tables 3.7 and 3.8 can be used to construct cumulative distribution functions of net cash income levels associated with each of the two strategies. These are shown in Figure 3.3. The

Table 3.8 System Performance under Strategy 2

| Net Cash Income | Corn | | | Soybeans | | | | |
|-----------------------|------------------|--------------------|------------------|----------|------------------|--------------------|------------------|-------|
| | Acres Planted | Acres Harvested | Average Yield | Price | Acres Planted | Acres Harvested | Average Yield | Price |
| 22065.26 | 174.08 | 174.08 | 115.07 | 2.07 | 305.02 | 305.02 | 31.23 | 6.17 |
| 13922.67 | 170.00 | 170.00 | 95.39 | 2.19 | 300.00 | 300.00 | 30.93 | 5.67 |
| -8576.60 | 144.76 | 115.57 | 70.54 | 2.14 | 335.24 | 335.24 | 21.12 | 6.20 |
| 20003.53 | 170.00 | 170.00 | 101.19 | 2.07 | 300.00 | 300.00 | 29.63 | 6.68 |
| 1644.65 | 155.36 | 83.61 | 60.25 | 2.30 | 321.64 | 304.96 | 28.91 | 6.19 |
| 12896.61 | 170.00 | 170.00 | 89.48 | 2.13 | 300.00 | 300.00 | 29.25 | 6.35 |
| 10057.10 | 157.21 | 157.21 | 90.85 | 2.07 | 322.79 | 322.79 | 28.25 | 6.20 |
| 12344.61 | 170.00 | 170.00 | 91.18 | 2.22 | 300.00 | 300.00 | 28.13 | 6.26 |
| -13618.91 | 180.00 | 170.00 | 91.10 | 1.87 | 300.00 | 300.00 | 26.05 | 4.55 |
| 34106.76 | 170.00 | 170.00 | 100.48 | 2.03 | 300.00 | 300.00 | 31.62 | 7.51 |
| 29832.77 | 170.00 | 170.00 | 95.45 | 2.31 | 300.00 | 300.00 | 28.92 | 7.63 |
| 29765.59 | 180.00 | 180.00 | 92.41 | 2.31 | 300.00 | 300.00 | 30.97 | 7.35 |
| -12271.70 | 139.34 | 77.55 | 51.50 | 1.83 | 341.66 | 341.66 | 25.08 | 5.09 |
| -14227.45 | 170.00 | 170.00 | 47.26 | 2.39 | 300.00 | 300.00 | 19.63 | 6.82 |
| 7406.12 | 160.00 | 170.00 | 75.53 | 2.07 | 282.38 | 282.38 | 25.41 | 7.77 |
| 5658.63 | 179.80 | 179.80 | 88.01 | 2.20 | 300.20 | 300.20 | 26.11 | 6.07 |
| -1402.99 | 170.00 | 170.00 | 93.96 | 1.91 | 300.00 | 300.00 | 27.62 | 5.61 |
| 13248.14 | 170.00 | 170.00 | 103.69 | 2.22 | 300.00 | 300.00 | 29.61 | 5.54 |
| 30765.34 | 170.00 | 170.00 | 106.61 | 2.20 | 300.00 | 300.00 | 31.19 | 7.09 |
| 26572.30 | 167.78 | 167.78 | 95.83 | 2.32 | 312.62 | 312.62 | 29.51 | 7.15 |

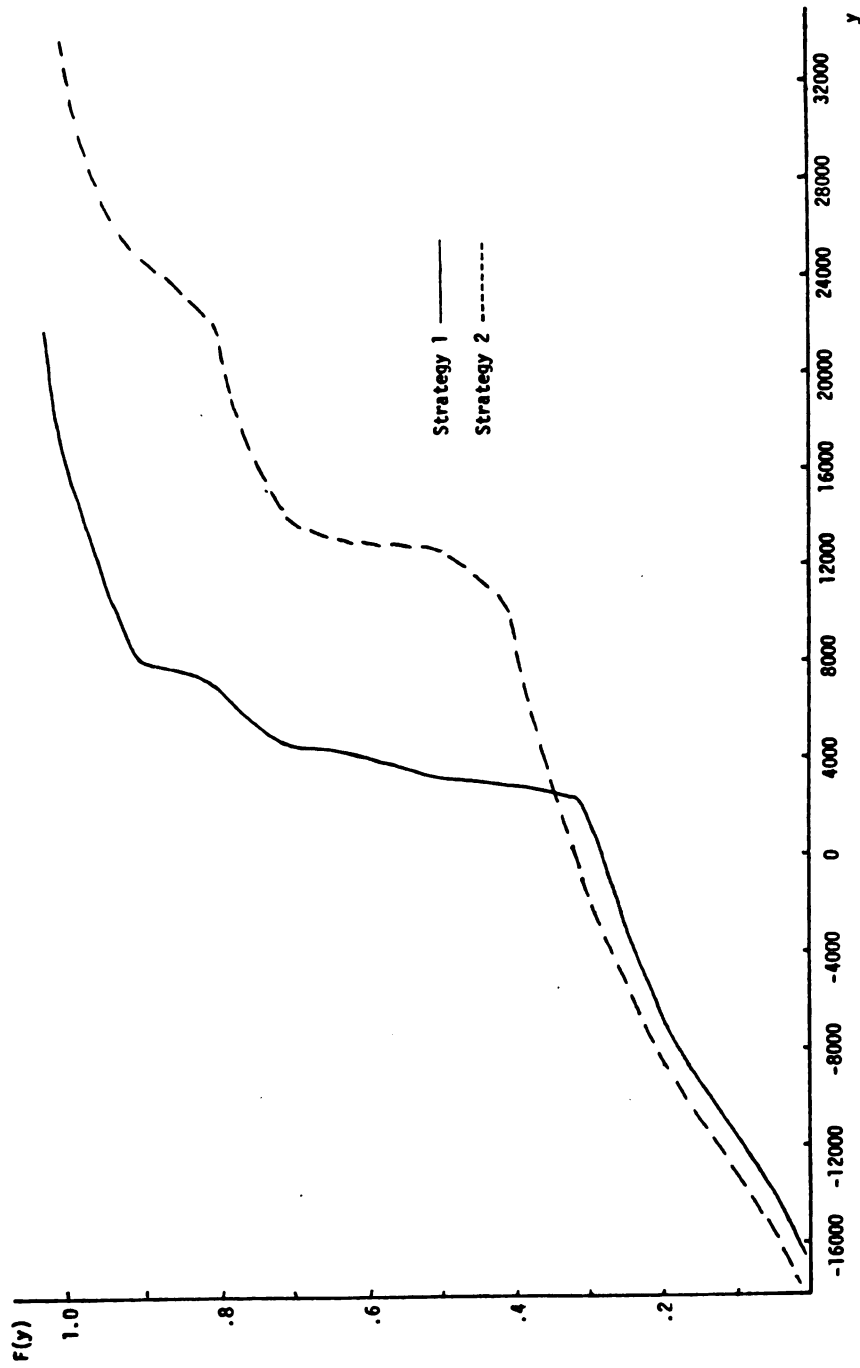


Figure 3.3 Cumulative Distribution Functions for Two Alternative Strategies

procedure for ordering alternative choices which is introduced in the next chapter requires that such a function be constructed for each strategy considered.

CHAPTER IV

THE MEASUREMENT OF DECISION MAKER PREFERENCES

4.1 Introduction

Choices made under uncertainty are affected by decision maker preferences for alternative outcomes as well as by subjective assessments of probability distributions of system outputs. When confronted with the choice between participation in two uncertain activities for which all possible outcomes and their probabilities are specified exactly, one decision maker may choose the first alternative while another may choose the second. This divergence in behavior cannot be attributed to a difference in subjective probabilities, since all relevant probabilities are specified prior to the time when a choice must be made. Rather, it must be attributed to a difference in the preferences of the two decision makers. Preferences, like assessments of probabilities in situations less highly structured than this example, are personal in nature, and some determination must be made of them in any applied decision analysis. This chapter examines procedures for eliciting information on decision maker preferences and techniques for combining this information with subjective probability assessments to identify preferred choices.

A decision maker's preferences can be represented quantitatively by a utility function,

$$U = u(y)$$

4.1

This is simply a relationship between the outcome of a choice as represented by a vector of system output levels, y , and an index of its desirability, U . It is a relationship which assigns values to alternative situations or conditions. When combined with a decision rule, such as utility maximization, a utility function becomes the basis for the identification of a preferred course of action.

If the system output levels associated with a particular strategy can be known with certainty, and if the decision maker's utility function is also known, calculation of the utility level of this choice is a relatively simple, direct matter. In uncertain decision situations, however, levels of system outputs realized under a specified strategy cannot be known exactly at the time when a choice is made, and the associated level of utility cannot be determined directly. In such situations the expected utility hypothesis provides a way of assigning assessments of value to alternative choices. First proposed in the eighteenth century by mathematician Daniel Bernoulli to explain the gambling behavior of some decision makers in uncertain situations, and derived more formally nearly 200 years later by Ramsey (1931) and by von Neumann and Morgenstern (1944), the expected utility hypothesis states that for any decision maker whose preferences are complete, transitive, continuous, and independent:

- (1) An ordinal utility function, $u(y)$, can be constructed such that $u(y)$ is defined for all system output levels and $u(y_1) > u(y_2)$ if the outcome defined by y_1 is preferred to that defined by y_2 .
- (2) The utility of an uncertain prospect is equal to the expected utility of its possible outcomes.

- (3) The scale of $u(y)$ is arbitrary up to a positive linear transformation--i.e., rankings of the utility function $V(y) = a + bu(y)$ are identical to those according to $u(y)$ if a and b are constants and $b > 0$.¹

The second result is the key one, since it is the basis for the commonly used decision rule which states that a decision maker's preferred choice is that which maximizes his expected utility.

If the expected utility hypothesis is to be applied in a practical context, the decision maker's utility function must be represented accurately enough to serve as a reliable aid in the identification of a preferred course of action. As is true in the determination of subjective probability distributions, however, the degree of accuracy sought in the measurement of preferences is dependent largely upon the characteristics of the decision problem under consideration. In some instances a very precise measure of preferences may be required, while in others nearly all feasible alternatives can be eliminated from consideration on the basis of only an approximate measure of preferences.

In the remaining sections of this chapter several alternative approaches to the measurement and representation of decision maker preferences will be examined along with the evaluative criteria used in conjunction with each type of measurement to order action choices.²

¹See Fishburn (1970) and Hirshleifer (1970) for more extensive discussions of the derivation of the expected utility hypothesis and for a complete explanation of the axioms which underlie it.

²Though the importance of preference measures based on more than one system output variable is recognized, the difficulty and cost of determining such measures preclude their use in most practical decision situations. The discussion below focuses entirely, then, on the measurement of preferences for outcomes which are adequately described by a single system output variable. See Keeney and Raiffa (1977) for an excellent discussion of the measurement of multidimensional preference relationships.

Each of the approaches considered is based on the expected utility hypothesis, but each requires a different level of precision in the measurement of preferences. Techniques for the representation of decision maker preferences and for the identification of preferred choices based on the use of single valued utility functions are reviewed first. The concept of an efficiency criterion, which allows the partial ordering of possible alternatives on the basis of relatively unrestrictive assumptions about decision maker preferences, is then introduced and several commonly used criteria are examined. A more recently developed efficiency criterion, stochastic dominance with respect to a function (Meyer, 1977a) is then described, and a procedure for determining the interval measurements of decision maker preferences required for the application of this criterion is presented. This new measurement technique, developed as part of this study, permits the construction of a representation of decision maker preferences which is only as precise as the decision problem under consideration requires. Results of an empirical test of this procedure, which are also presented below, demonstrate that it is both accurate and flexible. In the final section of the chapter, the incorporation of information on preferences into the sample problem discussed in the preceding two chapters is examined.

4.2 The Use of Single Valued Utility Functions to Represent Decision Maker Preferences

Perhaps the most direct approach to the measurement of preferences is to actually derive the decision maker's utility function. This requires that a number of points in the utility function be determined by

direct elicitation. A curve is then fitted through these points, and that curve is said to be the decision maker's utility function.

A utility function is a highly structured representation of a decision maker's preferences. Like probabilities, however, preferences are often not clearly formulated in the mind of the decision maker. Therefore, the interview procedures used to elicit information on preferences should be designed both to clarify and to structure the decision maker's assessments of value. Several procedures have been developed for the elicitation of information on preferences. The most commonly used are reviewed in Officer and Halter (1968) and Anderson, Dillon, and Hardaker (1977). Each procedure requires that a series of choices be made between pairs of uncertain alternatives or between certain and uncertain alternatives. If these choices are properly structured, each should reveal enough information about the decision maker's preferences to determine one point on his utility function.

Once a set of data points has been elicited, a curve is fitted through them to obtain an explicit relationship between levels of utility and all relevant levels of the system output variable. In choosing a functional form, a number of factors should be considered. Goodness of fit is, of course, important, since the estimated utility function should conform as closely as possible to the information obtained in the elicitation interview. Ease of estimation and the tractability of a function in the calculation of expected utilities should also be considered. Polynomial specifications are, perhaps, the most commonly used in empirical work, but a number of other alternative forms have also been prepared (Lin and Chang, 1978).

Having estimated a decision maker's utility function, action choices can be ordered by calculating the expected utility of each. If the outcome associated with a particular management strategy, v^* , is described by a single discrete system output variable, y , the expected utility of the strategy, $EU(y|v^*)$, is given by the expression:

$$EU(y|v^*) = \sum_{i=1}^n f(y_i|v^*)u(y_i) \quad 4.2$$

where $f(y_i|v^*)$ is the probability of the i^{th} possible outcome under strategy v^* and $u(y_i)$ is the utility of that outcome. When the system output variable, y , is continuous, the expected utility is defined by the expression:

$$EU(y|v^*) = \int_{-\infty}^{\infty} f(y|v^*)u(y)dy \quad 4.3$$

where $f(y|v^*)$ is the probability density function of y under strategy v^* . Consider, for example, the case in which the decision maker's utility function is of the form

$$u(y) = \ln(y) \quad 4.4$$

and the probability distributions associated with two strategies, A and B, are as specified in Table 4.1. Since the system output variable is discrete, equation 4.2 can be used to calculate the expected utilities for these two alternatives. That for strategy A is

$$\begin{aligned} EU_A &= .5 \ln(500) + .1 \ln(1000) + .1 \ln(1500) + .3 \ln(2000) \quad 4.5 \\ &= 6.81 \end{aligned}$$

while that for strategy B is

$$\begin{aligned} EU_B &= .2 \ln(500) + .6 \ln(1000) + .1 \ln(1500) + .1 \ln(2000) \quad 4.6 \\ &= 6.88 \end{aligned}$$

For this decision maker, then, strategy A is preferred to strategy B.

Table 4.1 Probability Distribution Associated with
Two Alternative Action Strategies

| System Output Level | Probability | |
|------------------------|-------------|------------|
| | Strategy A | Strategy B |
| 500 | .5 | .2 |
| 1000 | .1 | .6 |
| 1500 | .1 | .1 |
| 2000 | .3 | .1 |

Representation of a decision maker's preferences with a single-valued utility function has several serious shortcomings in an applied decision analysis. With regard to preference measurement procedures, the hypothetical choices posed in the elicitation interviews are, in general, less complex and less interesting than those actually facing a decision maker. As a result, it may be difficult to hold the full attention of the respondent through a series of similar questions. Furthermore, it can be argued that, because the types of choices made during the elicitation interview bear little resemblance to those made in real life, the value of the interview itself as a learning process whereby the respondent can gain a better understanding of how he makes decisions is limited. Other problems encountered in the elicitation of the information required to construct single valued utility functions are discussed in Officer and Halter (1968).

Still more serious problems arise as a result of the way empirically estimated utility functions are generally used in a decision analysis. Once a set of data points from a decision maker's utility function has been elicited, a curve is fitted through its elements. Rare indeed is the case in which the fit is perfect so that the parameter values of

the utility function can be known with certainty. Even if the fit were perfect, shortcomings of the preference elicitation procedures make it likely that the data points themselves include measurement errors. Therefore, an empirically estimated utility function cannot be considered to be an exact representation of decision maker preferences. Despite the possible sources of imprecision, however, a utility function, once estimated, is usually treated as though it were an exact representation of preferences when alternative choices are ordered, and any absolute difference in the expected utilities associated with two choices is taken as a clear indication that one is preferred to the other. If a utility function does not accurately reflect a decision maker's actual preferences, this can result in the recommendation of a choice which is not actually the preferred choice of the decision maker. When empirically estimated single valued utility functions are used to order alternative choices, then, there is a high likelihood that errors of this sort will be made.

4.3 Efficiency Criteria and the Representation of Decision Maker Preferences

The difficulties associated with the use of single valued utility functions to order choices in a practical context have been the impetus for the development of several efficiency criteria which overcome some of the shortcomings identified above. An efficiency criterion is a preference relationship which provides a partial ordering of feasible action choices for decision makers whose preferences conform to certain rather general specifications. As such, an efficiency criterion can be used to eliminate some feasible choices from consideration without

requiring detailed information about the decision maker's preferences. In many instances, the use of such a criterion may greatly reduce the number of alternatives to be considered. If enough alternatives can be eliminated, it may be possible for a final choice to be made on the basis of direct comparisons of the distributions of outcomes associated with each of the remaining alternatives.

First and second degree stochastic dominance are among the simplest and most commonly used efficiency criteria. Both were formulated independently by Hadar and Russell (1969) and Hanoch and Levy (1969). First degree stochastic dominance holds for all decision makers who prefer more of the system output to less--i.e., for all decision makers having positive marginal utility with respect to the system output variable. An alternative for which the associated distribution of the system output variable is described by the cumulative distribution function $F(y)$ is preferred to a second alternative with associated cumulative distribution $G(y)$ by the criterion of first degree stochastic dominance of

$$F(y) \leq G(y) \quad 4.7$$

for all possible levels of y and if the inequality in 4.7 is a strict inequality for at least some value of y . In Figure 4.1, for example, $F(y)$ dominates $G(y)$ by this criterion, since it is always below and to the right. Neither $F(y)$ nor $G(y)$ can be ordered with respect to $H(y)$ according to this criterion, since both are strictly greater than $H(y)$ for some system output levels.

While first degree stochastic dominance holds, in effect, for all decision makers, second degree stochastic dominance places an

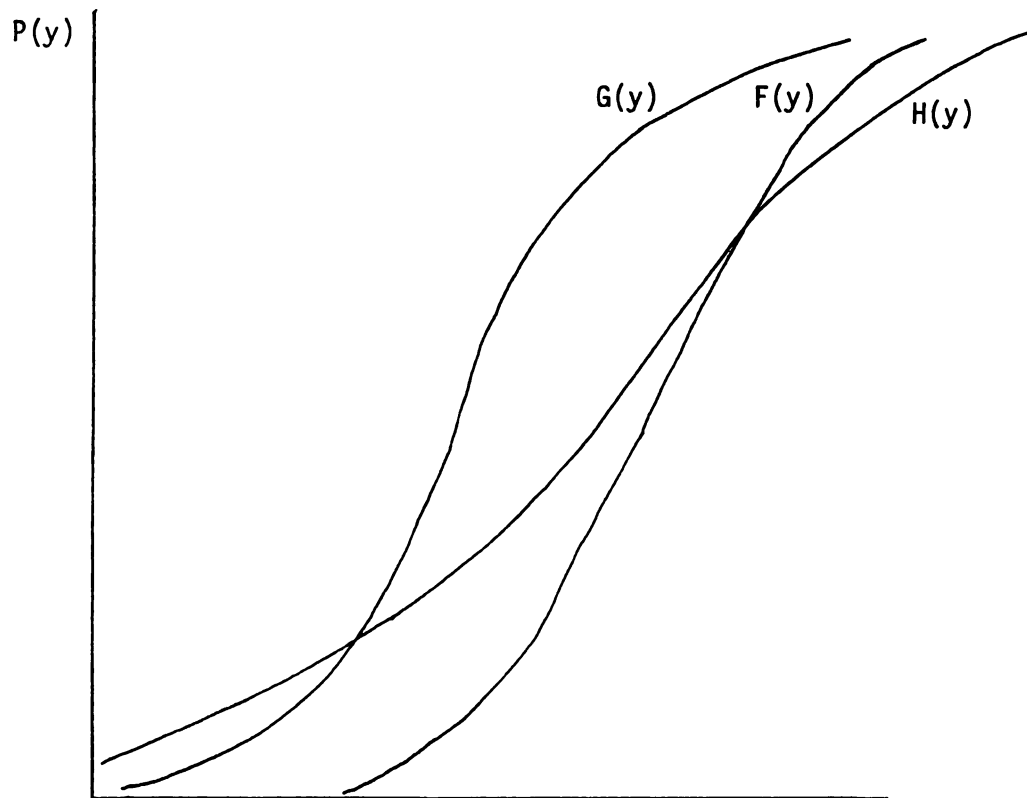


Figure 4.1 Illustrations of First and Second Degree Stochastic Dominance

additional restriction on preferences. It requires that the marginal utility of the system output variable be both positive and decreasing-- i.e., it requires that the decision maker's utility function be concave. Given two alternatives having system output distributions defined by the cumulative distribution functions $F(y)$ and $G(y)$, respectively, the first alternative is preferred to the second under the criterion of second degree stochastic dominance if

$$\int_{-\infty}^y F(x)dx \leq \int_{-\infty}^y G(x)dx \quad 4.8$$

for all possible values of y and if the inequality in 4.8 is a strict inequality for at least some value of y . In effect, this means that the first alternative dominates the second if the area under cumulative $F(y)$ is always less than or equal to that under $G(y)$. In Figure 4.1, for example, $F(y)$ dominates both $G(y)$ and $H(y)$ by this criterion, since the area under this cumulative is less than that under either of the others at all values of y . $G(y)$ and $H(y)$ cannot be ordered by this criterion, however, since the area under $H(y)$ is at times less than that under $G(y)$ and vice versa.

Other efficiency criteria depend on additional restrictions in the maker's preferences or in the nature of the probability distributions of system outputs. The mean-variance efficiency criterion (Markowitz, 1959) is simply a special case of second degree stochastic dominance in which all probability distributions are normal. Third degree stochastic dominance (Whitmore, 1970) is similar to first and second degree stochastic dominance, but it requires the additional assumption that the decision maker's utility function have a positive third derivative with respect to the system output variable.

Once a particular criterion with its associated restrictions on preferences has been specified, an ordering of any two alternatives can be made strictly on the basis of properties of the two associated probability distributions of the system output variable. Under such an ordering, one alternative will dominate the other, or the criterion will not be able to order the two alternatives and both will be considered efficient. If one alternative does dominate the other, it is unanimously preferred by the class of decision makers for whom the criterion applies. By making a series of pair-wise comparisons of all alternatives under consideration and eliminating from consideration any alternative which has been dominated, an efficient set of choices can be determined for any finite set of alternatives. This set will contain the preferred choice of any member of the class of decision makers for whom the criterion applies.

The use of an efficiency criterion to order alternative choices is, in many respects, preferable to the use of a single valued utility function. No direct measurements of preferences need be made. Rather, relatively easily accepted restrictions are simply imposed on the decision maker's preferences. Unfortunately, however, none of the efficiency criteria mentioned above is a particularly discriminating evaluative tool. In an application of second degree stochastic dominance by Anderson (1975), for example, twenty of forty-eight randomly generated farm plans were in the efficient set. Furthermore, though the restrictions on preferences required by most efficiency criteria do not appear to be unduly strict, they often run counter to empirical evidence. Again focusing attention on second degree stochastic

dominance, despite the fact that strong theoretical arguments have been made for the near universality of concave utility functions (Arrow, 1971), the weight of empirical evidence indicates that decision makers do at times exhibit increasing marginal utility (Officer and Halter, 1968; Conklin, Baquet, and Halter, 1977).

While the concept of an efficiency criterion is an attractive one, then, efficiency criteria have not proved to be useful tools in practice. There is a need for efficiency criteria which are both more flexible and more discriminating than those described above. Furthermore, there is a need for techniques for obtaining measures of decision maker preferences which, though less precise than those used to construct a single-valued utility function, facilitate the empirical determination of whether or not a particular efficiency criterion adequately represents the preferences of a decision maker. In the sections which follow, a more powerful efficiency criterion, stochastic dominance with respect to a function (Meyer, 1977a), is introduced, and a method for measuring decision maker preferences designed to be used in conjunction with this criterion is presented.

4.4 Stochastic Dominance with Respect to a Function

Stochastic dominance with respect to a function is an evaluative criterion which orders uncertain action choices for classes of decision makers defined by specified lower and upper bounds, $r_1(y)$ and $r_2(y)$, on the absolute risk aversion function. The absolute risk aversion function (Arrow, 1971; Pratt, 1964), $r(y)$, is defined by the expression:

$$r(y) = -u''(y)/u'(y) \quad 4.9$$

where $u'(y)$ and $u''(y)$ are the first and second derivatives of a von Neumann-Morgenstern utility function $u(y)$. In the most abstract terms, values of the absolute risk aversion function are simply local measures of the degree of concavity or convexity exhibited by a decision maker's utility function. Since $u'(y)$ is assumed to be positive, a positive value of $r(y)$ implies a negative value of $u''(y)$ which in turn implies a concave utility function. Similarly, the utility function is convex at y if $r(y)$ is negative. As such, the absolute risk aversion also serves as a local indicator of the extent to which a decision maker is risk averse or risk loving. Following Arrow's (1971) definition, an individual is risk averse (loving) if, from a position of uncertainty, he is unwilling (willing) to take a bet which is actuarially fair (unfair).¹ Concavity of the utility function and risk aversion are synonymous under this definition, and both are implied by a positive value of $r(y)$. A negative value of $r(y)$ implies both local convexity of the utility function and risk loving behavior. Perhaps the most important property of the absolute risk aversion function, however, is that it is a unique measure of preferences, while a utility function is

¹Arrow's definition of risk aversion has been the source of some confusion, since risk aversion and risk preference have often been equated with an aversion to and a love for gambling. Unless some measure of the degree of gambling associated with a particular choice is identified as a system output and included as an argument in a decision maker's utility function, however, his choices are, by the omission of this factor, assumed to be unaffected by the degree of gambling involved. Arrow's concept of risk aversion refers only to the characteristics of a utility function with a single argument. As Friedman and Savage (1948) demonstrate, such a utility function can be used to explain why gambling has utility or disutility in certain situations without requiring that preferences for gambling per se be measured.

unique only to a positive linear transformation.¹ In effect, then, upper and lower bounds on a decision maker's absolute risk aversion function define an interval measurement in his preferences. Stochastic dominance with respect to a function orders choices on the basis of such a measurement.

The major advantage of this criterion is that it imposes no restrictions on the width or shape of the relevant region of risk aversion shape. The interval measurement can be as precise or imprecise as is deemed necessary for a particular decision analysis. Negative as well as positive levels of absolute risk aversion can lie within the risk aversion interval at some or all levels of system output. Less flexible efficiency criteria such as first and second degree stochastic dominance can be viewed as special cases of this more general criterion. The requirement under first degree stochastic dominance that the decision maker have positive marginal utility places no restrictions on the decision maker's absolute risk aversion function--i.e., $r_1(y) = -\infty$ and $r_2(y) = \infty$ for all possible values of y . The requirement under second degree stochastic dominance that marginal utility be decreasing as well as positive, on the other hand, implies that $r_1(y) = 0$ and $r_2(y) = \infty$ for all values of y .

More formally stated, stochastic dominance with respect to a function is a criterion which establishes necessary and sufficient

¹ Because a utility function is unique only to a positive linear transformation, $u(y)$ and $u^*(y) = a + bu(y)$, $b > 0$ are strategically equivalent, though perhaps highly dissimilar, utility functions. The absolute risk aversion functions of these two utility functions are identical, however:

$$r(y) = -u''(y)/u'(y)$$

conditions for the distribution of system outputs defined by the cumulative distribution function $F(y)$ to be preferred to that defined by the cumulative distribution function $G(y)$ by all agents whose absolute risk aversion functions lie everywhere between lower and upper bounds $r_1(y)$ and $r_2(y)$. As developed by Meyer (1977a), the solution procedure requires the identification of a utility function $u_0(y)$ which minimizes

$$\int_0^1 [G(y)-F(y)]u'(y)dy \quad 4.9$$

subject to the constraint

$$r_1(y) \leq -u''(y)/u'(y) \leq r_2(y) \quad y \in [0, 1] \quad 4.10$$

The expression in equation 4.9 is equal to the difference between the expected utilities of system output distributions $F(y)$ and $G(y)$.² If, for a given class of decision makers, the minimum of this difference is positive, $F(y)$ is unanimously preferred to $G(y)$. If the minimum is zero, it is possible for an agent in the relevant class of decision makers to be indifferent between the two alternatives and they cannot be ordered. Should the minimum be negative, $F(y)$ cannot be said to be unanimously preferred to $G(y)$. In this case, the expression

$$\int_0^1 [F(y)-G(y)]u'(y)dy \quad 4.11$$

¹The range of system outputs is normalized so that all values of y fall on the bounded interval $[0, 1]$.

²This can be demonstrated in the following manner. Let $f(y)$ and $g(y)$ be the probability density functions associated with $F(y)$ and $G(y)$

$$\int_0^1 f(y)u(y)dy - \int_0^1 g(y)u(y)dy = \int_0^1 [f(y)-g(y)]u(y)dy$$

is the difference between the expected utilities associated with the two distributions. Integrating by parts,

$$\int_0^1 [f(y)-g(y)]u(y)dy = [F(y)-G(y)]u(y) \Big|_0^1 - \int_0^1 [F(y)-G(y)]u'(y)dy = \int_0^1 [G(y)-F(y)]u'(y)dy$$

since $[F(0)-G(0)]$ and $[F(1)-G(1)]$ are both equal to zero.

must then be minimized subject to 4.10 to determine whether $G(y)$ is unanimously preferred to $F(y)$. It should be noted that a complete ordering is not ensured by the criterion. It is possible for the minimum of both (4.9) and (4.11) to be negative, which implies that neither distribution is unanimously preferred by the class of decision makers being considered.

Meyer uses optional control techniques outlined in Arrow and Kurz (1970) to derive the necessary and sufficient conditions for the solution of this problem. These conditions do not represent a closed form solution. Rather, they define a rule for determining the absolute risk aversion function of the utility function which minimizes 4.9--a rule can be applied if the relatively unrestrictive assumption that $[G(y)-F(y)]$ changes a sign finite number of times over the interval $[0, 1]$ is met. The following theorem (Meyer, 1977a, p. 333) is the basis for that rule:

Theorem An optimal control $-u''_0(y)/u'_0(y)$ which minimizes

$$\int_0^1 [G(y)-F(y)]u'(y)dy \text{ subject to } r_1(y) \leq [u''_0(y)/u'_0(y)] \leq r_2(y)$$

and $u'_0(0) = 1$ is given by:

$$\frac{-u''_0(y)}{u'_0(y)} = \begin{cases} r_1(y) & \text{if } \int_y^1 [G(x)-F(x)]u'_0(x)dx < 0 \\ r_2(y) & \text{if } \int_y^1 [G(x)-F(x)]u'_0(x)dx \geq 0 \end{cases}$$

This theorem implies that the value of the absolute risk aversion function which minimizes the difference in the expected utilities associated with $F(y)$ and $G(y)$ is determined at any point y^* by the sign of the objective function integrated from y^* forward to 1 using the optimal control. Furthermore, it implies that the value of the absolute risk aversion function is always either $r_1(y)$ or $r_2(y)$.

Application of the rule requires that the solution procedure work from back to front as is commonly done in dynamic programming. A simple example should help to illustrate how this can be done. Consider the two cumulative distribution functions shown in Figure 4.2. Neither dominates the other by first or second degree stochastic dominance. Let the lower and upper bounds on the absolute risk aversion functions both be constants-- $r_1 = .001$ and $r_2 = .002$.¹ The utility function associated with each of these can be shown to be of the negative exponential form (Pratt, 1964), so that

$$u_i(y) = -e^{-r_i y} \quad i=1, 2, \quad 4.12$$

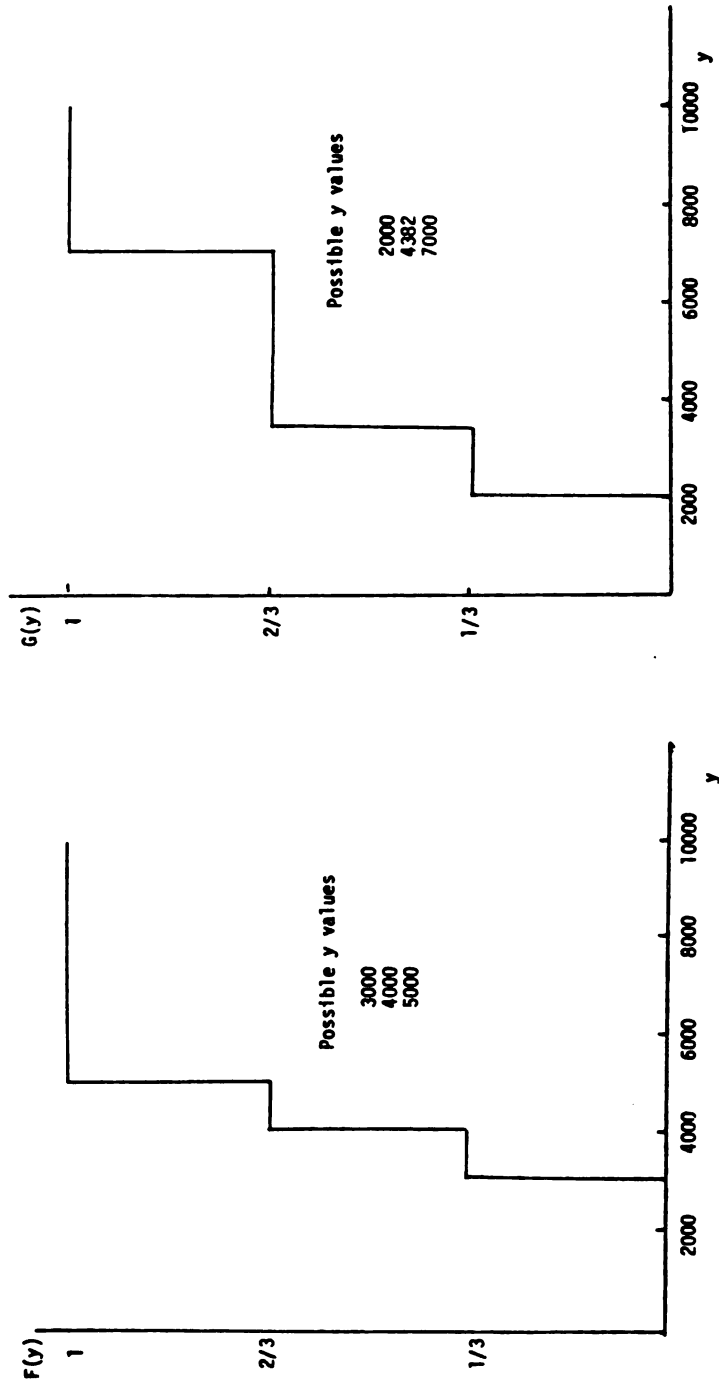
where r_i is the upper or lower bound absolute risk aversion level. The function $[G(y)-F(y)]$ is graphed in Figure 4.3. Between $y=5000$ and $y=7000$ its value is negative, and above $y=7000$ its value is zero. According to the theorem above, this implies that for values greater than 5000 $r_i(y) = .001$ is the optimal control. Calculating the value of the objective function from $y=5000$ and upward,

$$\begin{aligned} \int_{5000}^{\infty} [G(y)-F(y)]u'_0(y)dy & \quad 4.13 \\ &= \int_{5000}^{7000} (-1/3)(.001)e^{-.001y}dy \\ &= -.00194 \end{aligned}$$

we see that it is negative.² The solution rule indicates that the optimal control remains at $r(y)=.001$. The procedure continues back

¹There is no requirement that they be constant. This assumption is made here to facilitate calculation.

²The value of the integral is zero over regions where $[G(y)-F(y)]=0$.

Figure 4.2 Cumulative Distribution Functions $F(y)$ and $G(y)$

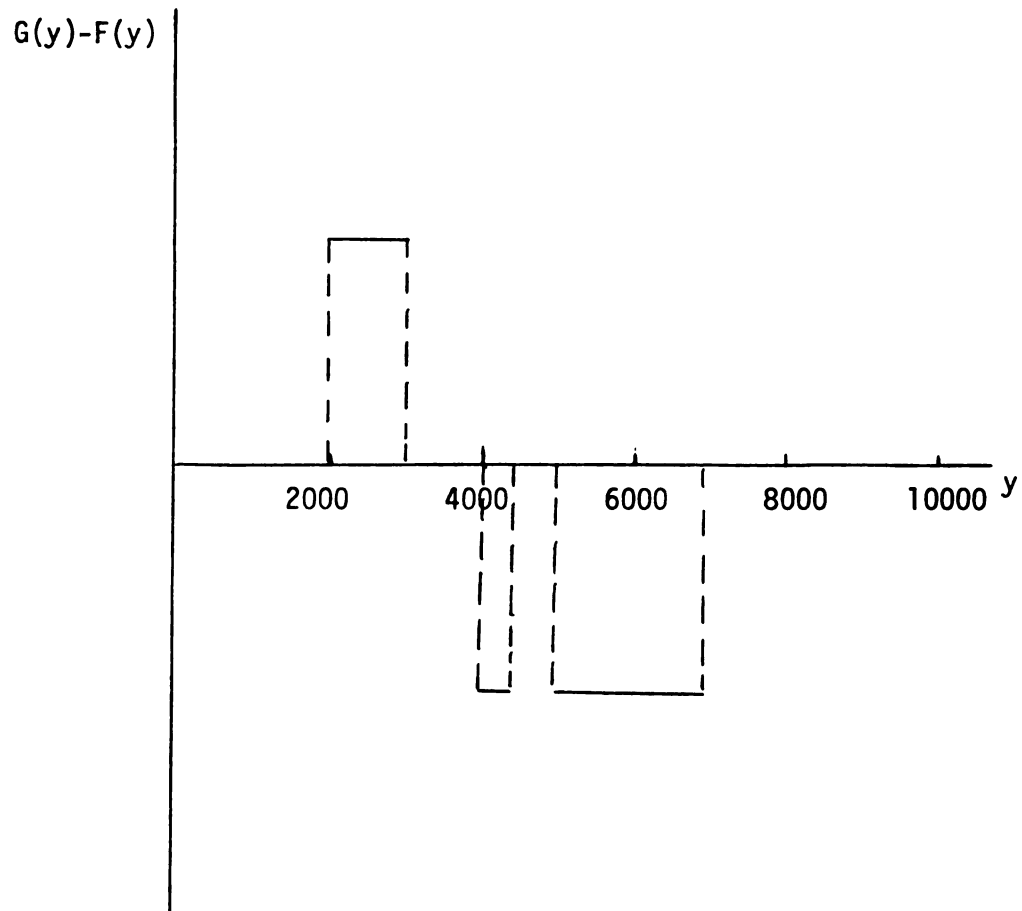


Figure 4.3 A Graph of the Function $[G(y)-F(y)]$

until the point where $y = 4000$. At this point

$$\begin{aligned}
 & \int_{4000}^{\infty} [G(y) - F(y)] u'_0(y) dy & 4.14 \\
 & = \int_{4000}^{4382} (-1/3)(.001)e^{-.001y} dy - .00194 \\
 & = 0.
 \end{aligned}$$

Therefore, the optimal control switches to $\gamma = .002$ for values of y less than 4000. The procedure continues back with the same optimal control until

$$\begin{aligned}
 & \int_{-\infty}^{\infty} [G(y) - F(y)] u'_0(y) dy & 4.15 \\
 & = \int_{2000}^{3000} (1/3)(.002)e^{-.002y} dy \\
 & = .00528
 \end{aligned}$$

Since the value of the objective function is positive, distribution $F(y)$ is preferred to $G(y)$ by all decision makers whose absolute risk aversion functions lie everywhere between $r_1 = .001$ and $r_2 = .002$. The utility function which minimizes the objective function has an absolute risk aversion function such that:

$$\begin{aligned}
 r(y) = & \quad .002 \text{ when } y \leq 4000 & 4.16 \\
 & \quad .001 \text{ when } y > 4000
 \end{aligned}$$

Note that this utility function does not have constant absolute risk aversion.

Stochastic dominance with respect to a function is a remarkably flexible evaluative criterion which has considerable potential for use in the analysis of practical decision problems. Unlike other efficiency criteria, it does not require that fixed restrictions be imposed on the representation of the decision makers' preferences; and, because the bounds on absolute risk aversion can be as close or as far apart as desired, stochastic dominance with respect to a function can be used to order more choices than can be done with other criteria. Unlike a single valued

utility function, it does not require that an exact representation of the decision maker's preferences be specified. Furthermore, stochastic dominance with respect to a function is relatively easy to apply. A computer program which implements the solution procedure defined above has been developed by Meyer, and a modified version of that program is presented in Appendix B.

4.5 An Interval Approach to the Measurement of Decision Maker Preferences

Stochastic dominance with respect to a function is a powerful analytical tool. Before it can be used in an applied context, however, an operational procedure must be developed for the determination of lower and upper bounds on a decision maker's absolute risk aversion function. A technique for making such interval measurements of decision maker preferences is introduced in this section. This procedure uses information revealed by a series of choices between carefully selected distributions to establish lower and upper bounds on an individual's absolute risk aversion function. The degree of precision with which preferences are measured--i.e., the size of the interval between the lower and upper bound functions--can be specified directly in accordance with the characteristics of the problem under consideration. At one extreme the interval can be of infinite width, and at the other extreme it can converge to a single line.

The procedure for constructing interval measurements of decision maker preferences is based on the fact that under certain conditions a choice between two outcome distributions defined over a relatively narrow range of system output levels divides absolute risk aversion space over that range into two regions: one consistent with the choice and one

inconsistent with it. The level of absolute risk aversion at which the division is made depends solely on the two distributions--i.e., their properties define the two regions. The decision maker's preferences, as revealed by his ordering of the two distributions, however, determine into which of these two regions his level of absolute risk aversion is said to fall. By confronting the decision maker with a series of choices between carefully selected pairs of distributions, the region of absolute risk aversion space which is consistent with the decision maker's preferences can repeatedly be divided. With each choice a portion of that region is shown to be inconsistent with the decision maker's preferences, and the interval measurement for the level of absolute risk aversion is narrowed. The procedure continues until a desired level of accuracy is attained. Upper and lower limits for the level of absolute risk aversion are determined at a number of income levels. These values are used to estimate upper and lower limits for the absolute risk aversion function over the relevant range of incomes.

The validity of the statement that a choice between two distributions is, under certain conditions, the basis for a division of absolute risk aversion space into regions consistent and inconsistent with a decision maker's revealed preferences can be demonstrated using concepts developed by Meyer in "Second Degree Stochastic Dominance with Respect to a Function." In that paper Meyer (1977b, p. 483) proves the following theorem:

Theorem For cumulative distributions $F(y)$ and $G(y)$

$$\int_0^y [G(x)-F(x)]dk(x) \geq 0 \quad y \in [0, 1] \text{ and}$$

$$\int_0^1 [G(x)-F(x)]dk(x) = 0 \quad \text{only if}$$

$$\int_y^1 [G(x)-F(x)]dk(x) \leq 0 \quad y \in [0, 1]$$

The theorem states that $F(y)$ is preferred to $G(y)$ by all decision makers more risk averse than the utility function $k(y)$ and that decision makers having utility function $k(y)$ are indifferent between the two distributions only if $G(y)$ is preferred to $F(y)$ by decision makers less risk averse than $k(y)$.¹ The function $k(y)$, then, can be considered to be a boundary function, since it separates a class of decision makers who prefer $F(y)$ from a class who prefer $G(y)$.

If the distributions $F(y)$ and $G(y)$ are defined over a narrow range of system output levels and if the decision maker's absolute risk aversion function can be approximated by a constant value λ over that range, preference for $F(y)$ implies that λ is greater than or equal to the minimum value of the absolute risk aversion associated with $k(y)$. Otherwise, the decision maker would be less risk averse than $k(y)$ and his choice would be inconsistent with expected utility maximization. Preference for $G(y)$, on the other hand, implies that λ is less than or equal to the maximum value of the absolute risk aversion function associated with $k(y)$, since $F(y)$ is preferred by all decision makers more risk averse than $k(y)$. It

¹Using Pratt's definition of risk aversion in the large, a decision maker with utility function $u(y)$ is more risk averse than $k(y)$ if

$$\frac{-k''(y)}{k'(y)} \leq \frac{-u''(y)}{u'(y)} \quad \forall y,$$

while he is less risk averse than $k(y)$ if

$$\frac{-k''(y)}{k'(y)} \geq \frac{-u''(y)}{u'(y)} \quad \forall y$$

Meyer (1977b) shows that $F(y)$ is preferred to $G(y)$ by all decision makers more risk averse than $k(y)$ if

$$\int_0^y [G(x) - F(x)] dk(x) \geq 0 \quad \forall y \in [0, 1]$$

and if the inequality is strict for some value of y . He also shows that $G(y)$ is preferred to $F(y)$ by all decision makers less risk averse than $k(y)$ if

$$\int_y^1 [G(x) - F(x)] dk(x) \leq 0 \quad \forall y \in [0, 1]$$

and if the inequality is strict for some value of y .

should be noted that the assumption that a decision maker's absolute risk aversion function can be adequately approximated by a constant value over a narrow range of system output levels is critical here. The theorem stated above does not imply that decision makers who prefer $F(y)$ to $G(y)$ are more risk averse than $k(y)$; nor does it imply that decision makers who prefer $G(y)$ to $F(y)$ are less risk averse than $k(y)$. With the assumption of constant absolute risk aversion in the neighborhood of a given system output level, however, it can be inferred that decision makers who prefer $F(y)$ to $G(y)$ are not less risk averse than $k(y)$ and those who prefer $G(y)$ to $F(y)$ are not more risk averse than $k(y)$.

The properties of a utility function which serves as a boundary function between two distributions are dependent upon the two distributions.¹ By careful selection of distributions, a boundary function can be placed anywhere in risk aversion space. A series of questions can be devised, then, which allows the repeated reduction of region of risk aversion space consistent with the revealed preferences of a decision maker, thereby, narrowing the interval measurement of absolute risk aversion.

A simple example should help to illustrate how the procedure works. Let the boundary function for two distributions, $k_1(y)$, have an absolute risk aversion function which lies everywhere on the interval (λ_1, λ_2) , and let the first distribution be preferred by decision makers more risk averse than $k(y)$ while the second is preferred by those less risk averse than $k_1(y)$. In this case the decision maker prefers the first distribution. Given the assumption of constant absolute risk aversion over this range of system output levels, this implies that his level of

¹A boundary function does not exist for each pair of distributions. One would not exist, for example, if one distribution dominates the other by first degree stochastic dominance. Similarly, the existence of one boundary function does not preclude the existence of others.

absolute risk aversion between y_1 and y_2 lies everywhere above λ_1 , as is shown in part (a) of Figure 4.4. If choices between two additional pairs of distributions indicate first that the level is greater than λ_2 and second that it is less than λ_3 , it can be inferred that the decision maker's level of absolute risk aversion over this range of system outputs lies between λ_2 and λ_3 , as is shown in part (c) of Figure 4.4. With each choice, then, a portion of the region of absolute risk aversion space consistent with prior choices is shown to be inconsistent with the decision maker's preferences and the interval measurement of absolute risk aversion is narrowed. Choices are presented to the decision maker until a desired level of accuracy is attained. An interval measurement of a decision maker's absolute risk aversion function can be constructed over a much broader range of system outputs by making interval measurements in the neighborhood of several system output levels and connecting known portions of the upper and lower bound absolute risk aversion function with linear segments, as is done in Figure 4.5. In this case direct interval measurements have been made in the neighborhood of three system output levels: 3,000; 10,000; and 17,000.

4.6 Implementation of the Procedure

The discussion above describes an iterative approach to the construction of interval measurements of a decision maker's absolute risk aversion function. It does little, however, to answer the basic operational questions of how appropriate distributions can be selected and of how the boundary interval for any pair of distributions can be identified. The techniques used to implement the interval approach to

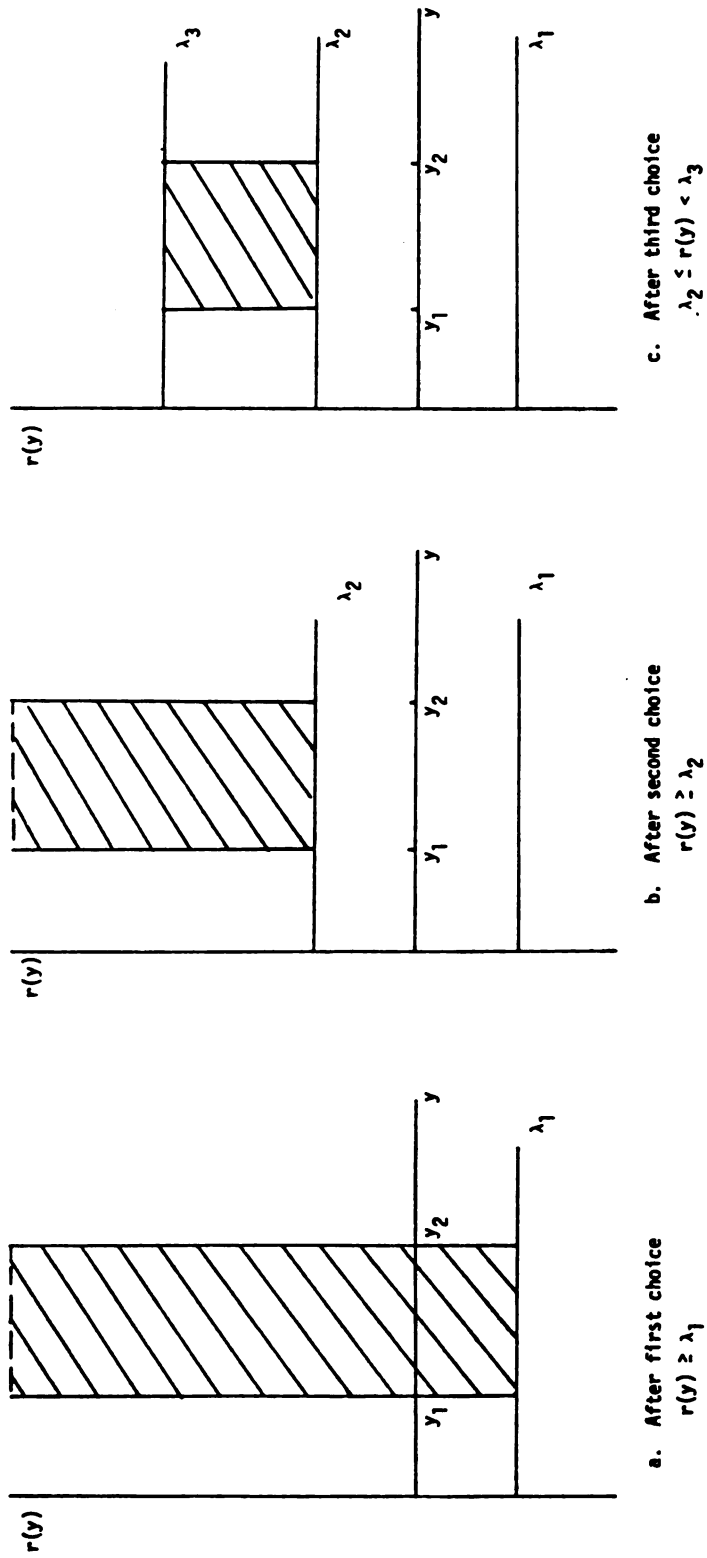


Figure 4.4 A Sequence of Interval Preference Measurements

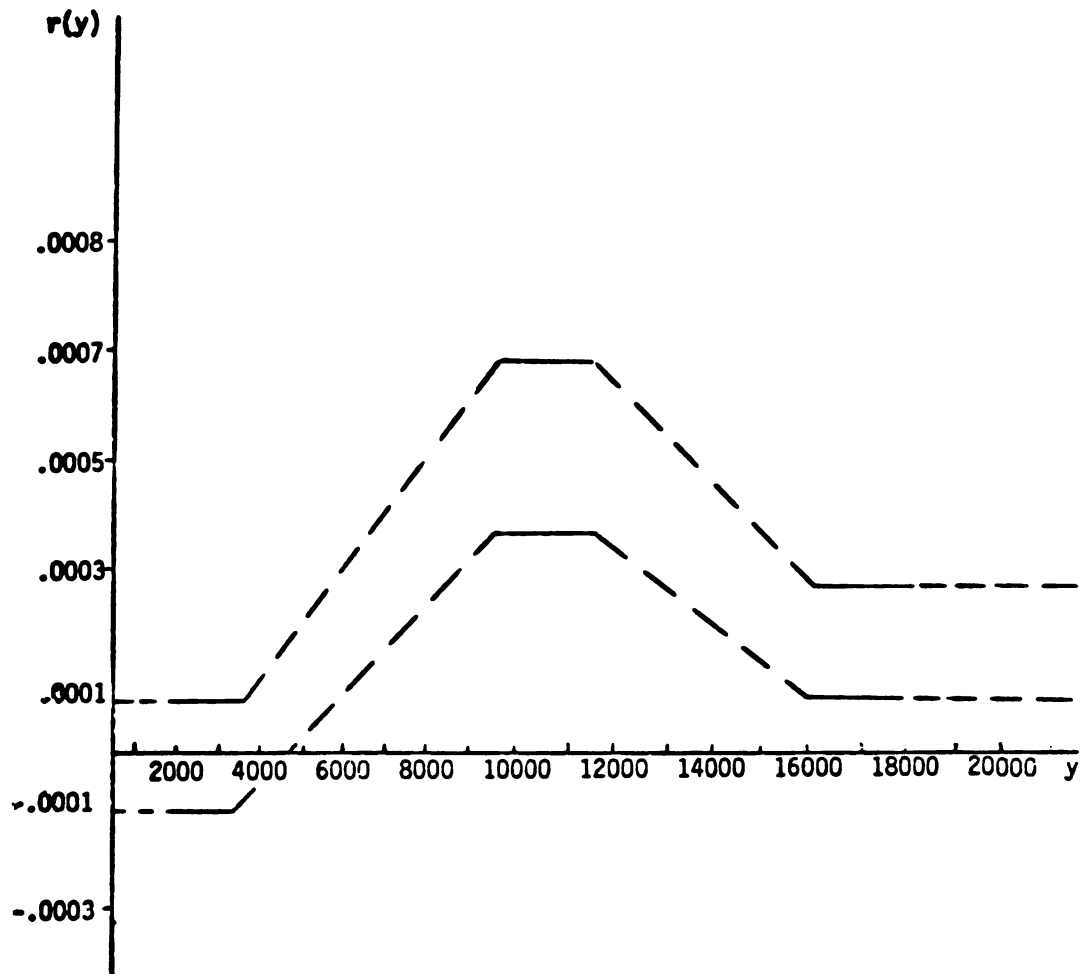


Figure 4.5 Upper and Lower Bound Absolute Risk Aversion Functions Based on Three Interval Measurements

the measurement of decision maker preferences are described briefly in this section. A more technical explanation of them is given in Appendix B.

The first step in the implementation of the procedure is to establish a measurement scale, which is defined by a number of reference levels for absolute risk aversion. In Figure 4.6, for example, four reference levels are specified: $-.0001$, $.0001$, $.0005$, and $.0010$. This scale or grid determines the accuracy with which absolute risk aversion can be measured. Any number of reference levels can be specified. The intervals between them can be as wide or narrow as is deemed necessary, and they need not all be of equal size. In many cases it may be desirable to put more fineness or detail in the measurement scale in regions of risk aversion space where it is believed a priori that the decision maker's level of absolute risk aversion is likely to fall.

Next, a set of distributions which will serve as the basis for the choices made by the decision maker must be constructed. These distributions should be defined over a relatively narrow range of system output levels, since the decision maker's level of absolute risk aversion is assumed to be constant over that range.¹ As described in Appendix B, they are constructed in a random manner by generating several hundred random numbers from a specified distribution and grouping them into sets of six observations each. Each is a distribution of outcomes, and each element is said to have a $1/6$ probability of occurrence. Only six elements are included in each distribution because more complex distributions might make decision makers' choices unduly

¹ Experience to date indicates that a range of five to ten percent the size of the entire range of system output levels over which preferences are to be measured is adequate.

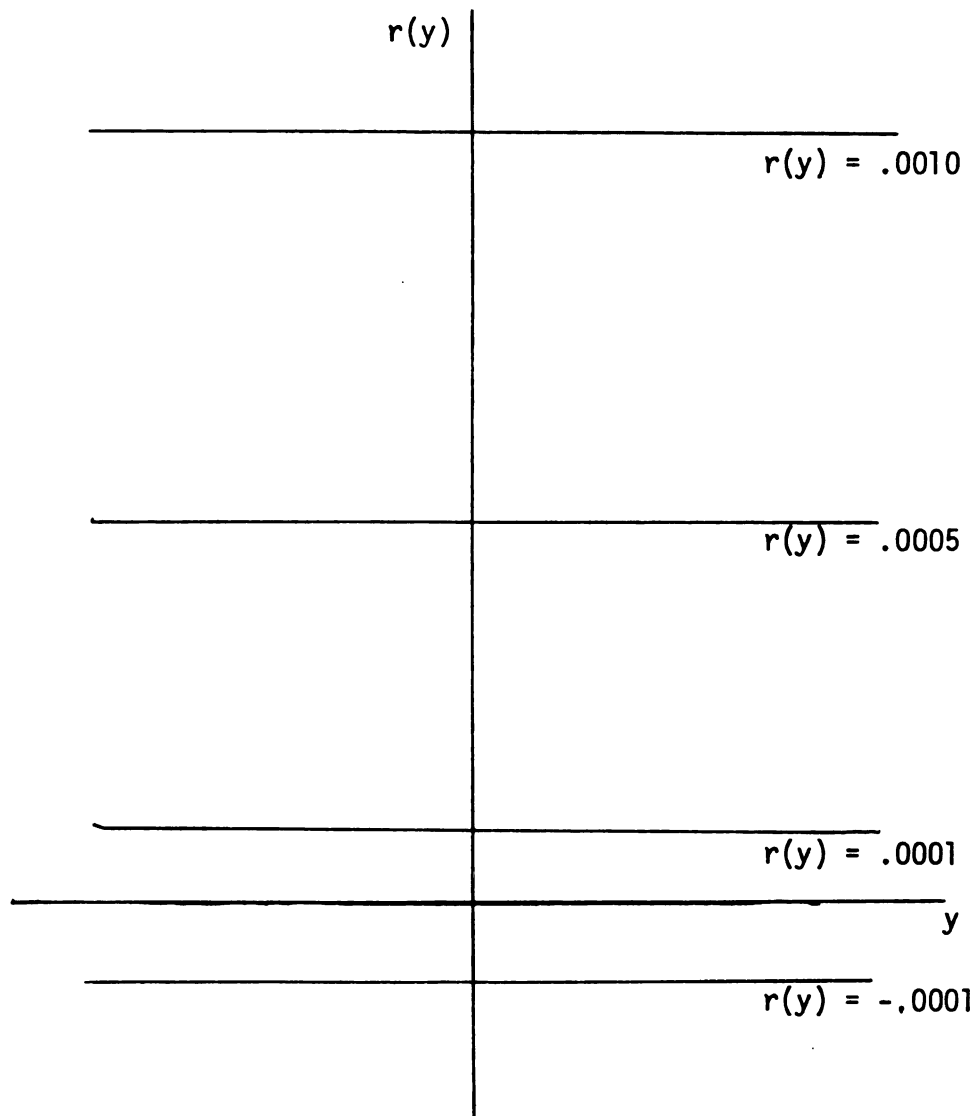


Figure 4.6 An Absolute Risk Aversion Measurement Scale

difficult. Distributions with fewer elements, on the other hand, may not be rich enough to make the choices interesting. The use of six element distributions also facilitates explanation of the choice situation to the decision maker, since the probability of any one element occurring can be equated directly to the probability of obtaining a specified number of dots on a single roll of a die. Three distributions constructed in this manner are shown in Table 4.2.

Once a measurement scale has been specified and sample distributions have been constructed, the interval on the measurement scale which contains the risk aversion function associated with the boundary function for each pair of distributions must be identified. The procedure by which this is done is described in detail in Appendix B. In essence, criteria identified by Meyer (1977b) are used to identify the highest reference level on the measurement scale, λ_1 , such that all decision makers less risk averse than λ_1 prefer one distribution and the lowest reference level in the measurement scale, λ_2 , such that all decision makers more risk averse than λ_2 prefer the other distribution. It follows that the absolute risk aversion function associated with the boundary function for the two distributions lies everywhere within the interval (λ_1, λ_2) , which is called a boundary interval. The boundary intervals for each pair of distributions from Table 4.2 are given in Table 4.3. In this case each pair has a different boundary interval. It should be noted that more narrow boundary intervals could have been identified if a more detailed measurement scale had been specified.

After the boundary interval has been identified for each pair of distributions, a series of questions is formulated. Each question asks

Table 4.2 Sample Distribution from a Normal Distribution
With $\mu = 3000$ and $\sigma = 1000$

| Distribution ^a | | |
|---------------------------|-----------------|----------------|
| 1 | 2 | 3 |
| 2100 | 1000 | 1750 |
| 2400 | 2050 | 1950 |
| 2550 | 2650 | 2500 |
| 3100 | 3800 | 2750 |
| 3250 | 3900 | 3950 |
| 3450 | 5200 | 4000 |
| $\mu = 2808$ | $\mu = 3100$ | $\mu = 2817$ |
| $\sigma = 488$ | $\sigma = 1370$ | $\sigma = 883$ |

^aThe elements of each distribution are rounded to the nearest 50 units.

Table 4.3 Boundary Intervals for Pairs
of Sample Distributions

| Distributions | Boundary Interval | Distribution Preferred Below Boundary Interval | Distribution Preferred Above Boundary Interval |
|---------------|-------------------|---|---|
| 1 VS 2 | (.0001,.0005) | 2 | 1 |
| 1 VS 3 | (-.0001,.0001) | 3 | 1 |
| 2 VS 3 | (.0005,.0010) | 2 | 3 |

the decision maker to indicate which of two selected distributions he prefers. His responses serve as the basis for the interval measurement of absolute risk aversion. Each question focuses on a particular interval of risk aversion space--an interval which corresponds to the boundary interval for the two distributions the decision maker is asked to rank. Let the first question in the example being developed here be:

Compare distributions 1 and 2 and indicate which
you prefer.

This question focuses on the interval $(.0001, .0005)$.¹ If distribution 1 is preferred, the information in Table 4.2 indicates that the decision maker's level of absolute risk aversion is greater than .0001. If distribution 2 is preferred, his level of absolute risk aversion is below .0005.

The choice of a second question will depend on the respondent's answer to the first. If the respondent prefers distribution 1, for example, it makes little sense to ask him to rank distributions 1 and 3. Such a question indicates whether his level of absolute risk aversion is greater than $-.0001$ or less than .0001. Since his level of absolute risk aversion is already said to be greater than .0001, this information would be of value only as a consistency check. A choice between distributions 2 and 3, on the other hand, would add to our knowledge. If distribution 2 is preferred, the decision maker's level of absolute risk aversion can be said to fall on the interval $(.0001, .0010)$, while if distribution 3 is preferred, the level must be greater than .0005.

¹It is advisable to focus the first question on a boundary interval at or near the middle of the measurement scale.

Because the order of questioning is contingent upon the decision maker's responses, the interview schedule takes a form similar to that of a programmed learning text, as is exemplified in Figure 4.7.

Clearly, the set of questions specified in this example does not lead to an accurate measure of a decision maker's preferences. By specifying a finer measurement scale, however, and by asking a series of three or four questions rather than only two, much more accurate measurements can be made. It should also be noted that these questions serve as the basis for a measurement of absolute risk aversion only for system output values between 1000 and 5000. Similar sets of questions must also be constructed to measure preferences in the neighborhood of several other system output levels so that upper and lower bound absolute risk aversion functions can be constructed for a wider range of possible system output levels.

Figure 4.5 above shows upper and lower bounds on an absolute risk aversion function based on interval measurements at three system output levels. Note that the slope of the absolute risk aversion function is not restricted. In this example it rises, then falls, and at lower levels of system output the measurement interval includes negative as well as positive values. When absolute risk aversion functions are derived from empirically estimated utility functions, on the other hand, their form is often severely limited by the functional form used to estimate the utility function. It should also be noted that the interval approach to the measurement of preferences also avoids another common problem encountered in the estimation of single-valued utility functions. Because all questions posed require a choice between two uncertain

1. Compare distributions 1 and 2 and circle the one you prefer.

1

2

If you prefer distribution 1 go to question 3,
otherwise go to question 2.

2. Compare distributions 1 and 3 and circle the one you prefer.

1

3

3. Compare distributions 2 and 3 and circle the one you prefer.

2

3

Figure 4.7 A Sample Questionnaire

prospects, biases due to preference for an aversion to gambling per se are eliminated.

One final step is required for the implementation of the interval approach to the measurement of preferences. Although Meyer's (1977a) analytical development of stochastic dominance with respect to a function depends only on absolute risk aversion functions, the computer program he has developed to implement the criterion requires that utility functions having absolute risk aversion functions corresponding to lower and upper bounds, $r_1(y)$ and $r_2(y)$, be specified by the user. Given the definition of the absolute risk aversion function,

$$r(y) = - \frac{u''(y)}{u'(y)} \quad 4.17$$

the following system of differential equations, which relates levels of absolute risk aversion to values of $u(y)$ and $u'(y)$ can be derived:

$$\frac{d}{dy} \begin{bmatrix} u(y) \\ u'(y) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -r(y) & 0 \end{bmatrix} \begin{bmatrix} u(y) \\ u'(y) \end{bmatrix} \quad 4.18$$

Once initial values of $u(y)$ and $u'(y)$ have been specified, recursive numerical integration techniques can be used to solve for $u(y)$ and $u'(y)$ at any level of system output.¹ Utility functions associated with the lower and upper absolute risk aversion functions are represented by a table look-up routine in the computer algorithm which implements the stochastic dominance with respect to a function criterion. Corresponding values of y and $u(y)$ determined by numerical integration serve as data points for each table look-up function.

¹The initial values of $u(y)$ and $u'(y)$ correspond to the arbitrary scale factors of a von Neumann-Morgenstern utility function. See Appendix B for a listing of the computer program which performs the numerical integration.

4.7 An Empirical Test

A simple experiment was designed and conducted to test the efficacy of the interval approach to the measurement of decision maker preferences. Three questionnaires were administered to a group of graduate research assistants from the Department of Agricultural Economics at Michigan State University. The first questionnaire employed the procedure described above to obtain an interval measurement of each subject's absolute risk aversion function. The second questionnaire was used to elicit information required for the construction of a single-valued utility function for each subject.¹ Finally, the third questionnaire asked the respondent to make a series of six choices between pairs of distributions, each distribution being comprised of six elements and each being defined on the interval over which preferences had been measured. Information from the first two questionnaires was used to predict the choices made by each respondent in the third questionnaire, and these predictions were compared to the actual responses. In this way the accuracy of each of the two approaches to the measurement of preferences was tested.

In evaluating each approach, two criteria were considered: the number of correct predictions and the number of choices for which a definite ordering was made. A prediction was said to be correct if the respondents' actual choice was not excluded from the efficient set of choices and incorrect if it was excluded. The preference measure having the highest proportion of correct predictions was said to be the more

¹Because it is the most commonly used and most easily implemented elicitation technique, the ELCE method (Anderson, Dillon, and Hardaker, 1977) was used to identify points on the decision maker's utility functions.

accurate according to this criterion. Concern with the proportion of correct predictions is analogous to concern with the probability of Type I error in a statistical test, the latter being the probability that a true statement will be judged to be false and be rejected. This measure of accuracy is not a good indicator of the relative discriminatory power of preference measurements based on these two approaches. The criterion of first degree stochastic dominance, which holds for all decision makers who prefer more of the system output to less, should never exclude a preferred choice from the efficient set and so should be perfectly accurate according to the criterion defined above. Often, however, it also fails to exclude many choices from the efficient set. A single-valued utility function, on the other hand, is the basis for a complete ordering of choices--i.e., it always leads to an efficient set having a single element. Therefore, the number of choices actually ordered was also considered. Concern with this measure of discriminatory power is analogous to concern with the Type II error associated with a statistical test, which is the probability that a false statement will be judged to be true and not rejected.

Clearly there are trade-offs between the accuracy and the discriminating power of a preference measurement. Unlike other measurement techniques and evaluative criteria, the combined use of interval preference measurements and stochastic dominance with respect to a function permits explicit consideration of these trade-offs. As the precision of the interval measurement increases, it becomes a more discriminating basis for the ordering of choices; but the probability of excluding preferred choices from the efficient set also increases.

Such trade-offs between accuracy and discriminatory power were also analyzed in the experimental test of the interval approach to the measurement of preferences. Direct interval measurements of absolute risk aversion were made at three levels of income--the relevant system output variable in this instance. These measurements were based on a sequence of four questions at each income level. By constructing interval measurements on the basis of information available at the end of each question, however, four preference measurements--each more precise than the one which preceded it--were made for each subject.

Nine of ten subjects correctly completed all three questionnaires. Since each subject made six choices on the third questionnaire, each preference measurement was used to predict a total of fifty-four choices. The results of the experiment are presented in Table 4.4. They show that there is a clear trade-off between accuracy and discriminatory power. First degree stochastic dominance and the single-valued utility function are at opposite extremes in this trade-off relationship, and the interval measurements are arrayed between the two. Several factors should be noted. With regard to the accuracy of the interval measurements, it falls at a relatively constant rate as the number of questions posed increases, but even at the higher levels of precision it exceeds that realized with the single-valued utility function. The discriminating power of the interval measurements, on the other hand, increases dramatically as the number of questions asked at each income level increases. First and second degree stochastic dominance, on the other hand, clearly do not discriminate well among the distributions which were the basis for the decision makers' choices.

Table 4.4 Performance Indicators for Alternative Preference Measures

| Performance Indicator | Interval Measurement | | | | Single Valued Utility Function | First Degree Stochastic Dominance | Second Degree Stochastic Dominance |
|---|----------------------|----|----|----|--------------------------------|-----------------------------------|------------------------------------|
| | Number of Questions | | | | | | |
| | 1 | 2 | 3 | 4 | | | |
| 1. Percent of choices predicted correctly | 98 | 88 | 78 | 72 | 65 | 100 | 98 |
| 2. Percent of choices ordered | 9 | 50 | 83 | 91 | 100 | 0 | 7 |

It should be noted that these results represent but one test of the interval approach to the measurement of decision maker preferences. Results presented in the next section and in Chapter V provide additional evidence of the power of this approach, but further experimentation is needed. It should also be noted that few attempts have been made to apply this technique. With more applied experience may come refinements that will improve the discriminatory power of preference measurements based on this approach without leading to increases in the probability of excluding the preferred choice from the efficient set.

4.8 An Application

In order to test the interval approach to the measurement of decision maker preferences in a more practical setting, questionnaires implementing the procedure were administered to seventeen farmers who were all participants in an extension workshop on cash grain marketing strategies. The questionnaires were viewed as an exercise in the workshop--an exercise designed to help individuals think systematically about how they make decisions. Interval measurements of absolute risk aversion were made in the neighborhood of four income levels: -\$3,000, \$7,000, \$17,000, and \$27,000. Each measurement was based on a series of three questions. In addition, the respondents were also asked to make a series of choices between distributions, as was done in the experiment described in the preceding section.

The farmers had little difficulty in completing the questionnaires, and they seemed to find the choices to be interesting. The range of responses was quite broad. Individuals within the sample of seventeen ranged from the extremely risk averse to the extremely risk loving.

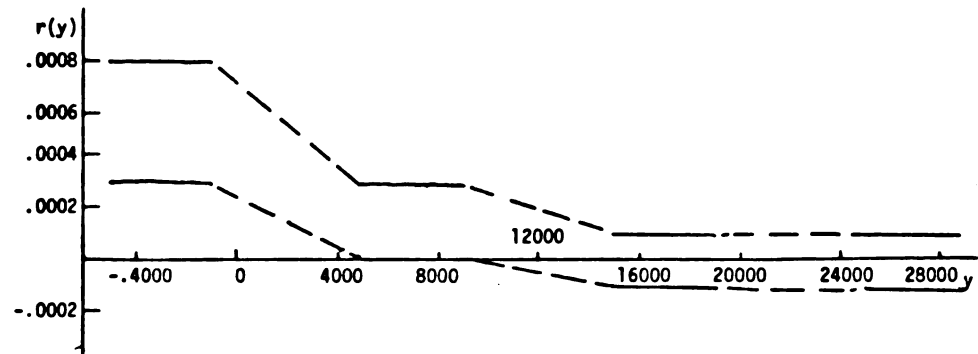
Several discernable patterns did emerge, however. Most decision makers exhibited increasing absolute risk aversion over the lower income levels and decreasing absolute risk aversion at higher levels. For most, the interval measurement of absolute risk aversion included negative values at some level of income. In fact, only four of the seventeen decision makers had lower level absolute risk aversion functions which were everywhere non-negative. This casts serious doubt upon the applicability of a criterion such as second degree stochastic dominance which is valid only for decision makers who are risk averse at all system output levels.

Choices made in the final section of the questionnaire were predicted remarkably well by the preference measures. Ninety-one out of 102 possible choices were predicted correctly for a success rate of .892. This compares quite favorably with that obtained in the more carefully controlled experiment with student subjects. To test the discriminatory power of the preference measures, that derived for each decision maker was used to order a set of thirty-three distributions, none of which was dominated by any other by the criterion of first degree stochastic dominance. The resultant efficient sets ranged in size from one to twenty-three, with the average size being 10.5. Clearly this represents a sizeable reduction in the size of the efficient set over that attained with first degree stochastic dominance. The second degree stochastic dominance efficient set for these thirty-three distributions had only five elements. It must be remembered, however, that this criterion is valid only for four of the seventeen farmers whose preferences were measured. For these four individuals the size of the efficient set averaged only 2.5.

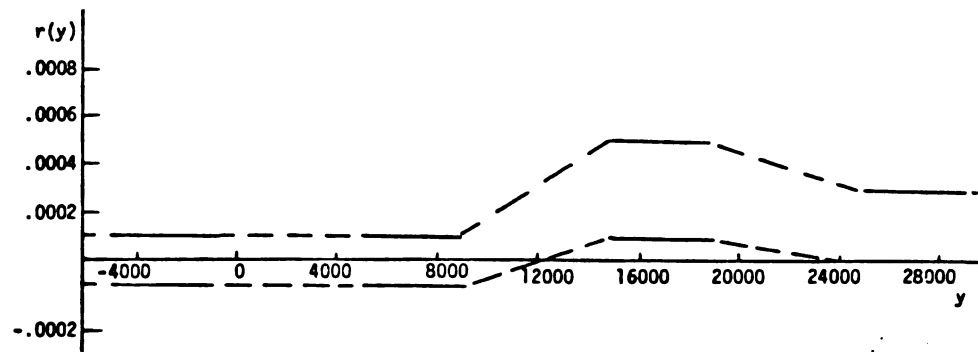
Upper and lower bound absolute risk aversion functions for three representative decision makers are shown in parts a, b, and c of Figure 4.8. The interval measurement for decision maker A declines across the relevant range of net income levels. That for decision maker B rises and then falls, while that for decision maker C is constant and then rises at the highest income level. Decision maker C is one of the four individuals in the sample who is everywhere risk averse.

The distributions of net income levels associated with each of the two strategies defined in Section 3.3 of Chapter III were ordered for each of these three decision makers using the criterion of stochastic dominance with respect to a function. It will be recalled that Strategy 1 calls for no land rental and for the planting of all acreage in soybeans. Strategy 2 calls for the rental of 240 acres and for a more balanced crop mix. Strategy 1 is preferred to Strategy 2 by decision maker A. Neither distribution dominates the other given the preferences of decision makers B and C.

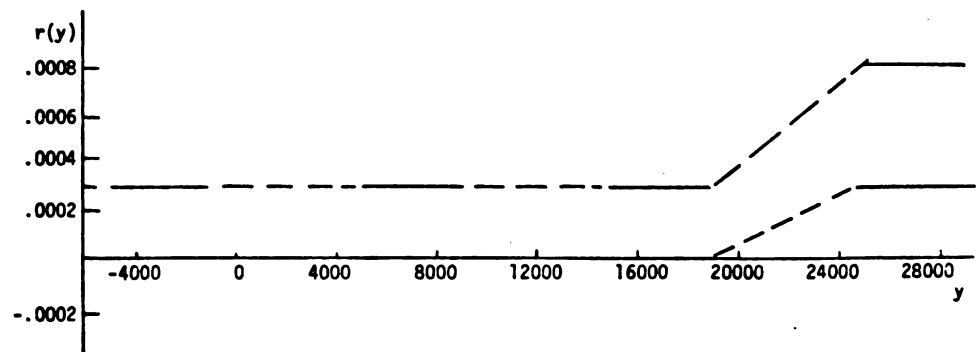
In themselves these results are not particularly interesting. They do demonstrate, however, that the interval measurements of decision maker preferences can lead to different efficient sets for different decision makers. They also show that remarkably dissimilar distributions may be included in an efficient set. The two strategies considered here bear little resemblance to each other. One could be called extremely cautious and the other moderately risky. For decision maker A it is clear that the cautious approach is preferred. For the other two decision makers the preferred choice is not so clear-cut. If these two strategies were but two of many being considered and if a single-valued



^aInterval Measurement of Decision Maker A



^bInterval Measurement for Decision Maker B



^cInterval Measurement for Decision Maker C

Figure 4.8 Interval Measurements of Absolute Risk Aversion for Three Decision Makers

utility function were being used to identify a preferred choice, one would have been eliminated from consideration. As the experimental results reported in the preceding section indicate, however, there is a relatively high probability that the preferred strategy would have been the one eliminated.

CHAPTER V

COMPUTATIONAL PROCEDURES FOR THE IDENTIFICATION
OF PREFERRED CHOICES

5.1 Introduction

Techniques developed in the preceding chapters for the determination and representation of subjective probability distributions and for the measurement of decision maker preferences provide the information required to order any two specified action choices. In most decision situations, however, a large if not infinite range of choices is open to the decision maker. As a result, some systematic technique for the identification and evaluation of a large number of possible strategies is needed in many applied decision analyses. Such a technique should be flexible enough to be applicable in a wide range of decision situations without requiring that important simplifying assumptions be made concerning preferences, probabilities, and the nature of the problem itself. It should serve as an aid in solving practical problems, without forcing the decision maker to alter his conceptualization of the problem at hand.

In this chapter a computational procedure designed to meet these needs is formulated, and its implementation is discussed. This procedure integrates concepts and operational techniques related to problem formulation, the determination of subjective probability distributions, and the measurement of decision maker preferences developed in the preceding chapters. It is a decision aid which is both powerful and highly flexible.

In subsequent sections of this chapter, existing computational procedures for the identification of preferred choices are first reviewed, and their strengths and weaknesses are identified. The procedure developed for this study is then introduced and described in detail. Finally, the procedure is applied to the sample problem discussed in the preceding three chapters.

5.2 A Review of Existing Computational Procedures

Mathematical programming models are commonly used in the analysis of complex decision problems when the assumption of perfect knowledge can reasonably be made. They are analytically elegant, computationally efficient, and easily adapted for use in a wide range of decision situations. A number of difficulties are encountered, however, when mathematical programming models are employed in the analysis of decisions made under uncertainty--difficulties related to problem formulation, to the determination of the probability distributions for system output variables associated with alternative strategies, and to the representation of decision maker preferences. Despite such difficulties, mathematical programming techniques are the basis for the most commonly used computational procedures for the identification of preferred choices under uncertainty.

Quadratic programming (Markowitz, 1959; Freund, 1956) is perhaps the most familiar and the most widely accepted mathematical programming technique for the analysis of decisions made under uncertainty. Conceptually it is an attractive tool because it is so closely linked with mean-variance analysis, which has been the basis for a wide

range of theoretical developments, and because when used in a parametric programming mode it can, under certain conditions, be used to identify an efficient set of strategies which includes the preferred choice of any risk averse decision maker.¹ With respect to practical considerations, it is an attractive technique because the formulation of quadratic programming problems is only slightly more complex than that of a standard linear programming problem and because quadratic programming packages are available in most computer systems. The usefulness of quadratic programming in an applied decision analysis is severely limited, however, by a number of other factors.

With regard to problem formulation, standard quadratic programming models require that input-output relationships be linear and additive and that all controllable system input levels be perfectly divisible. In many practical decision situations these assumptions simply do not correspond closely with reality. Equally serious are the limitations imposed by the standard quadratic programming model on the definition of a management strategy. Decisions are analyzed as though they were inflexible, though, as was noted in Chapter II, one of the most important characteristics of choices under uncertainty is that they are often adaptive in nature. Finally, though computer codes which implement quadratic programming algorithms are readily available, many limit the size of the problem which can be considered.

With regard to the determination of probability distributions, quadratic programming requires that they be determined analytically,

¹As was noted in Section 4.3 of Chapter IV, the criterion of mean-variance efficiency is a special case of second degree stochastic dominance. The mean-variance efficient set is identified by parametric quadratic programming.

which may greatly limit the types of stochastic factors which can be considered in an analysis and may also limit the complexity of the model used to represent a particular stochastic process. Furthermore, within a quadratic programming framework, the distributions of exogenous system inputs and of system output variables are described only by means, variances, and covariances. Implicitly, then, all distributions are assumed to be normal. When this assumption does not hold, quadratic programming may eliminate from consideration the preferred choices of some decision makers (Tsiang, 1972; Robison and King, 1978).

Finally, with respect to the representation of decision maker preferences, quadratic programming requires that the decision maker's utility function be of the quadratic or negative exponential form if a single preferred choice is to be identified. When parametric quadratic programming is used to identify a mean-variance efficient set, on the other hand, the efficient set holds only for risk averse decision makers. It is possible that a decision maker's preferences cannot be adequately represented under any of these assumptions.

In response to some of the shortcomings, several linear programming alternatives to quadratic programming have been proposed. These include the game theoretic (McInerney, 1969; Hazell, 1970; Low, 1974) and focus-loss (Boussard and Petit, 1967) approaches and the MOTAD model developed by Hazell (1971). All these models can be solved using standard linear programming algorithms, and none requires that stochastic returns be normally distributed. They impose fewer limitations on the size of problems which can be considered and all are more amenable to the relaxation of restrictions on linear constraints and production processes and on divisibilities of choice variables through the use of

separable programming and mixed integer programming techniques. Each does have serious shortcomings, however. Covariance between returns for different activities is ignored in both the MOTAD and focus-loss models. This can lead to a serious misrepresentation of the distribution of outcomes associated with any particular action choice. More important, however, the links between the decision criteria used in these approaches and the expected utility hypothesis are much weaker than is the case with quadratic programming. The safety-first behavioral assumption implied by the focus-loss approach and employed in many applications of the game theoretic model is especially difficult to reconcile with the axioms underlying the expected utility hypothesis.

The Risk Efficient Monte Carlo Programming (REMP) model developed by Anderson (1975, 1976) is in nearly all respects a more attractive alternative to quadratic programming as a decision aid in practical situations. The REMP model employs Monte Carlo programming techniques (Donaldson and Webster, 1968) to construct a large number of feasible management strategies in a random fashion. The distribution of total net returns associated with each strategy under consideration is determined analytically under the assumption that distributions of net returns for each activity and distributions of total net returns for each strategy are members of the beta family. Covariance between returns for each activity is considered explicitly. Under this procedure, a cumulative distribution function is constructed for total net returns associated with each strategy and the criterion of second degree stochastic dominance is used to identify an efficient set of choices. The REMP model allows considerable flexibility in the

representation of probability distributions, since the beta distribution can take a variety of forms. The model also places few restrictions on decision maker preferences. The criterion of second degree stochastic dominance requires only that the decision maker be risk averse at all levels of system output. With regard to problem formulation, the REMP model also has several distinct advantages. Most notably, the use of Monte Carlo programming allows, at least partially, the relaxation of restrictive assumptions of linearity, additivity, and perfect divisibility required by most mathematical programming models. Despite these advantages, the usefulness of the REMP model is greatly limited by the fact that second degree stochastic dominance is not a very discriminatory criterion. In many instances the size of the efficient set identified with the REMP model is so large that the task of selecting a single preferred strategy may be prohibitively difficult.

None of these alternatives to linear programming resolves the problems associated with the incorporation of explicit consideration of flexibility into the decision analysis. Recursive programming techniques (Day, 1963; Heidhues, 1966) do permit the consideration of such factors within a linear or quadratic programming framework. The behavioral constraints, feedback rules, and repeated optimization which characterize this approach are used to model the process of planning, decision making, and action. The constraints and rules of thumb which drive the model through this process--the desired output of a decision analysis--must be determined exogenously, however. The value of the recursive programming as an aid to decision makers, then, is limited. Stochastic programming (Cocks, 1968; Rae, 1971) is conceptually a more

attractive approach to the resolution of this problem. Adaptive decision strategies are determined endogenously under this procedure, which can be used with a standard linear or quadratic programming algorithm. If the problem under consideration is even moderately complex, however, the size of the input-output matrix quickly expands to an unmanageable level, so the usefulness of this approach as an applied decision analysis is also limited. Linear decision models, as developed by Holt, Modigliani, Muth, and Simon (1966), represent a third alternative. Under this approach, which has been further refined by Zellner (1971), Chow (1973), and McRae (1975), dynamic programming techniques are used to determine analytical solutions for a special class of optimal control problems. Having assumed quadratic cost (utility) functions and linear relations between state and control variables, linear decision rules are derived which can be used to determine optimal levels for control variables on the basis of forecasts of key factors. The parameters of these rules remain invariant for as long as the system design parameters are unchanged. Models of this sort are subject to the same limitations imposed on preferences and probability distribution under quadratic programming. Furthermore, the complexity of such models and the high cost performing the initial analysis needed to determine optimal decision rules for a particular process limit applicability of this approach to situations where similar decision are made repeatedly.

The models described above also fail to completely resolve the problem that complex stochastic processes can often not be adequately represented in a quadratic programming framework. Of particular

importance is the fact that the impact of uncertainty concerning fixed resource availability levels cannot be satisfactorily analyzed in the models discussed above. Such factors can be important in some situations. In the decision problem discussed in the preceding three chapters, for example, the number of days available for fieldwork during any planting or harvest period is highly uncertain and has a major impact on crop yields and net income levels. Consideration of this type of uncertainty is commonly incorporated into programming models using chance constrained programming techniques (Charnes and Cooper, 1959). Though relatively simple to implement, however, this approach does not permit explicit consideration of the cost associated with violating a constraint and so does not actually facilitate the explicit determination of the impact of stochastic fixed resource levels on the distribution of system output levels realized under any particular management strategy.

Finally, none of these models, as specified, permits the representation of preferences by an empirically determined interval measurement of absolute risk aversion in the ordering of choices by the criterion of stochastic dominance with respect to a function, the advantages of which were demonstrated in Chapter IV. It should be noted, however, that stochastic dominance with respect to a function can be rather easily incorporated into the REMP model, as will be demonstrated in the next section.

5.3 A Generalized Procedure for the Identification of Preferred Choices Under Uncertainty

While each of the decision models discussed above is attractive in light of at least one theoretical or practical consideration and

while each may be appropriate for use in some decision situations, none can be said to be generally applicable in a wide range of practical contexts. Furthermore, many rather common types of decision problems cannot be adequately analyzed with any of the procedures discussed above. A more general approach is needed--one which permits greater flexibility with respect to problem formulation, the determination of probability distributions, and the representation of decision maker preferences without sacrificing the power of decision theory based on the expected utility hypothesis. Such an approach is presented in this section.

The generalized procedure for the identification of preferred choices described here is in many respects an extension of Anderson's (1975, 1976) REMP model. Feasible strategies are generated using a modified form of the Monte Carlo programming model developed by Donaldson and Webster (1968), which also serves as a basis for Anderson's model. Under the more generalized procedure, however, a strategy may be defined by specific controllable system input levels, by a set of adaptive decision rules, or by some combination of the two. Distributions of system output levels associated with particular strategies are not determined analytically. Rather, they are determined by simulating system performance in a number of sample states of nature, using techniques described in Chapter III. This facilitates the consideration of the impact of a wide range of stochastic exogenous system inputs and permits greater flexibility in the representation of complex stochastic processes. Finally, strategies are evaluated using interval measurements of decision maker preferences and the evaluative criterion of stochastic dominance with respect to a function.

Like the REMP model this procedure is an iterative one. A large number of strategies are generated and evaluated sequentially. The determination of a truly optimal choice is not ensured. If a sufficiently large number of plans is examined, however, it is reasonable that the efficient set will contain a nearly optimal choice for a decision maker.

Because of its similarity to the REMP model, this generalized procedure for the identification of preferred choices under uncertainty can be called the generalized risk efficient Monte Carlo programming model (GREMP). Interrelationships among the three major processes within the model--strategy generation, system output distribution determination, and evaluation--are illustrated by the flow chart in Figure 5.1. Each of these processes will be discussed in greater detail in the remainder of this section.

5.3.1 Generation of a Feasible Management Strategy

At the outset of each iteration of the GREMP model a management strategy is constructed. As defined in Chapter II, a management strategy is a set of controllable system input levels, a set of feedback control rules for determining controllable system input levels over the duration of the planning horizon, or some combination of the two. The nature of the problem under consideration determines the types of choices which must be made, and the nature of the decision situation determines the range of choices open to the decision maker.

Regardless of how decisions are defined, the presence of constraints makes some management strategies impossible. Limits on available resources may restrict the set of admissible values for some controllable system inputs. Similarly, some controllable system inputs may be

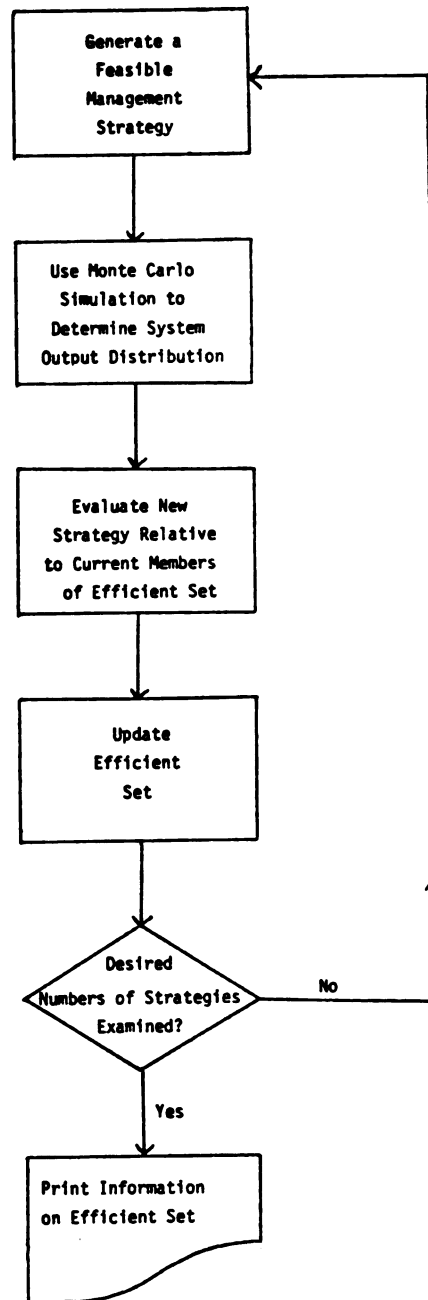


Figure 5.1 A Flow Chart of the GREMP Model

indivisible and so must take integer values. In addition, logic or common sense may dictate that a particular parameter of a feedback control rule should be positive or that one parameter should always be greater than another. Constraints may also be of a form which renders two activities mutually exclusive--e.g., the choice of a particular functional form for a feedback control rule precludes the use of an alternative form. Given the definition of a management strategy in a particular decision situation and the constraints on the range of available choices, some method of identifying feasible strategies for consideration in the decision analysis is needed. When the number of alternatives is small, each can be explicitly specified and evaluated. When the number of alternatives is large, Monte Carlo programming techniques can be a valuable tool for the identification of feasible strategies.

Monte Carlo programming is a search procedure which constructs sample management strategies at random from the set of feasible strategies. In determining the values for controllable system inputs and/or feedback control rule parameters, techniques similar to those introduced in the discussion of the simulation of stochastic processes are used. Monte Carlo programming is a remarkably flexible tool which can be relatively easily and inexpensively implemented and so is well suited for use in an applied decision analysis.

Monte Carlo programming techniques are explained in detail by Donaldson and Webster (1968). A technical discussion of their application is also presented in Appendix C of this study along with a listing of the computer program used to implement the GREMP model. In the context of current discussion an example is, perhaps, the most effective medium

for the explanation of the process by which Monte Carlo programming is used to construct feasible strategies. The management strategy specified in relation to the decision problem introduced in Chapter II will be the basis for this example. It will be recalled that this strategy is defined by three controllable system input levels and by one simple feedback control rule which has a single parameter. The three controllable system input variables are:

v_1 = number of acres rented

v_2 = number of acres planted in corn

v_3 = number of acres planted in soybeans

The feedback control rule is: "Regardless of specified crop acreage levels, soybeans will be planted on all unplanted acreage after v_4 (a parameter indicating a specific date)."

Four indivisible 80 acre tracts of land are available for rental. Therefore, the only admissible values for v_1 are 0, 80, 160, 240, and 320. Two types of constraints are imposed on crop acreage levels. First, the farmer states that if he grows a crop at all, he wishes to plant at least fifty acres of that crop, i.e.

$$v_1 = 0 \text{ or } v_i \geq 50 \quad i = 2, 3 \quad 5.1$$

Second, total crop acreage is restricted to that which is owned by the farmer, 240 acres, plus that which is rented, v_1 . The farmer wishes to plant all available acreage, if possible. Therefore:

$$v_2 + v_3 = 240 + v_1 \quad 5.2$$

With regard to the control rule parameter v_4 , three possible values will be considered: May 18, May 26, and June 3. As was noted in Chapter II, these are the ending dates of the last three possible corn planting

periods. Let $v_4 = 1$ if May 18 is chosen, 2 if May 26 is chosen, and 3 if June 3 is chosen.

The construction of a management strategy is a sequential process, and in some cases it may be necessary to set a value for one choice variable before levels of others can be established. In the example being discussed here, for example, a value of v_1 must be determined before crop acreage levels can be specified. Therefore, after control variables have been identified and constraints on them have been specified, the control variables must be classified according to the sequence in which they should be considered. In our example, land rental can be classified as a resource acquisition activity, and v_1 is the first variable for which a value is determined. Crop acreage levels, v_2 and v_3 , refer to resource using activities and are specified next. Finally, the value of the control rule of parameter, v_4 , is set.¹

In constructing a management strategy, the value of each controllable system input or feedback control rule parameter is treated as a random variable. In our example, v_1 and v_4 are clearly discrete random variables since each has only a few possible values. Unless it is desirable to assign greater probability weight to some particular value, both can be treated as discrete uniform random variables. In the case of land rental, then, there is probability of 1/5 that no land will be rented, 1/5 that 80 acres will be rented, 1/5 that 160 acres will be rented, 1/5 that 240 acres will be rented, and 1/5 that 320 acres will

¹The order in which control variable levels are specified depends, in general, on the characteristics of the problem under consideration and on computational convenience. All that is actually required in this example is that a value of v_1 be set before v_2 and v_3 are considered.

be rented. Similarly, a probability of $1/3$ can be assigned to each of the three possible values of the control rule parameter.

Acreage levels for each crop, v_2 and v_3 , can be treated as continuous or discrete random variables. Again a uniform distribution, be it discrete or continuous, should be used unless there are a priori reasons for assigning higher probabilities of selection to certain values or ranges of values. In this case acreage levels will be treated as discrete variables with possible values being integer multiples of ten lying on the interval between 50 acres and total available acreage, $240 + v_1$. The probability of any particular admissible value being selected is given by the expression:

$$P = \frac{1}{.10(240 - v_1 - 50) + 1} \quad 5.3$$

The denominator on the right hand side of 5.3 is a general expression for the number of possible values of v_2 or v_3 given a particular value of v_1 .

Actual construction of a management strategy begins with the generation of a sample observation from the distribution of v_1 . In this instance let the value of that observation be 240, which implies that 240 acres of land are to be rented under this strategy. Next, values of v_2 and v_3 , the crop acreage levels, must be determined. Again using a discrete uniform process generator, one of these two variables is selected for consideration. Let v_2 , the number of acres planted in corn, be the variable selected. A sample observation from the distribution of this variable is then generated. In this case let its value be 220 acres. Clearly the constraint on total acreage is not violated by this value, since v_3 is considered to be equal to zero until

it is assigned another value. This is, then, a feasible value for v_2 . The constraint on remaining available acreage is next updated to state that:

$$\begin{aligned} v_3 &= 240 + v_1 - v_2 & 5.4 \\ &= 240 + 240 - 220 = 260 \end{aligned}$$

and a value for v_3 is, in turn, established. It should be noted that had the value of v_2 been greater than 430, a value of v_3 less than 50 would have been implied. Because values less than 50 are not permitted for v_2 or v_3 , v_3 would have been set to zero and corn acreage would have been expanded to 480. It should also be noted that the process of control variable level determination is somewhat more complex when more than two resource-using variables enter into a single constraint. The procedures used in such cases are explained in Donaldson and Webster (1969) and in Appendix C, and the computer program developed for the implementation of the GREMP model is fully applicable to problems of this sort.

Once values for the three controllable systems input variables have been specified, all that remains in constructing a feasible management strategy is the determination of a value for the control rule parameter, v_4 . A sample observation from this variable's distribution is selected-- in this instance let its value be 2, which implies a date of May 26-- and the strategy is complete. It is:

v_1 = 240 acres of land rented

v_2 = 220 acres of corn to be planted

v_3 = 260 acres of soybeans to be planted

v_4 = May 26, the date after which all unplanted acreage is shifted to soybean production regardless of specified values of v_2 and v_3 .

This simple example demonstrates some of the features which make Monte Carlo programming such a flexible procedure. Choice variables can be continuous or discrete, and they need not be assigned values which correspond to strictly quantifiable entities. In addition, both upper and lower limits can be established for a choice variable once activated without necessarily forcing it into the management strategy at a non-zero level. Other features of Monte Carlo programming are outlined in Appendix C.

5.3.2 Determination of the Distribution of System Output Levels

Once a strategy has been constructed, the associated distribution of system output levels must be determined. This is done sequentially as each strategy is generated using the Monte Carlo simulation techniques described in Section 3.3 of Chapter III. System performance under a given strategy is simulated for a large number of sample states of nature, each defined by a sample vector from the joint probability distribution of relevant stochastic system input variables. In this way sample system output levels are determined, which can be used to construct a cumulative distribution function for the underlying distribution.

In the current example, net cash income is the system output variable which serves as the basis for the evaluation of alternative strategies. It will be recalled that crop prices, crop yields, and time available for fieldwork were judged to be the stochastic exogenous system inputs which have an important impact on system performance. Subjective probability distributions for all of these factors were

specified in Section 3.4 of Chapter III. Using techniques described in Appendix A, twenty sample vectors of levels for each of these stochastic system inputs were generated and read into the computer program which implements the GREMP model. A computer simulation model which determines the net cash income level realized under a specific management strategy in any given state of nature was also described in Section 3.4 of Chapter III. This model is incorporated into the larger GREMP model as a subroutine which is called after the generation of each alternative strategy. Using the stochastic system input data from the main program and the control variable levels for the new strategy, it calculates net cash income for each of the twenty states of nature. The twenty sample income levels associated with the management strategy defined in the preceding subsection are given in Figure 5.2.

5.3.3 The Evaluation of Alternative Strategies

Alternative management strategies are evaluated within the GREMP model by applying the criterion of stochastic dominance with respect to a function, with an interval measurement of decision maker preferences defining the relevant levels and upper bound absolute risk aversion functions. Evaluations are made sequentially as strategies are generated. If a particular strategy is not dominated by any current member of the efficient set, it, too, becomes a member of the efficient. In such instances the control variable values which define the new strategy and the associated set of sample system output levels are stored. If the new strategy is dominated by any member of the current efficient set, on the other hand, it is eliminated from further consideration. Similarly,

-17271.70
-15571.83
- 8576.60
- 1328.38
1644.65
5658.63
7705.34
10057.10
11737.75
13146.29
13676.88
15136.41
19763.05
22865.28
26578.39
28998.96
29472.42
29695.14
33951.82

$\mu = \$10,707.58$

$\sigma = \$15,464.33$

Figure 5.2 Sample Observations from the Distribution
of Outcomes Associated with Strategy One

members of the efficient set which are dominated by the new strategy are removed from the efficient set. Because the criterion of stochastic dominance with respect to a function is fully transitive, this procedure ensures that no member of the efficient set for the set of strategies already examined will be eliminated and that only information on actual members of the current efficient set will be saved.

Returning to the example being discussed, let the decision maker be the first of the three decision makers for whom actual interval measurements were described in Section 4.8 of Chapter IV. Both the lower and upper bounds of his risk aversion interval, it will be recalled, decrease monotonically over the range of income levels for which preferences were measured. Let the first strategy generated by the GREMP model be that defined above in Section 5.3.1. Being the first strategy considered, it automatically becomes a member of the efficient set.

Let the second strategy generated be:

v_1 = 0 acres rented

v_2 = 50 acres of corn to be planted

v_3 = 190 acres of soybeans to be planted

v_4 = May 26, the date after which all unplanted acreage is to be planted in soybeans

The set of net income levels associated with this strategy is given in Figure 5.3. Given this decision maker's preferences, this distribution dominates the distribution of income levels associated with the first strategy. After two iterations, then, the efficient set is comprised only of the second management strategy.

Let the third strategy generated by the GREMP model be defined by the following control variable levels:

-11482.39
- 9829.65
- 9463.32
- 5351.16
- 3452.86
- 606.25
731.79
2097.90
3043.55
3214.38
3545.55
3700.89
5900.30
6354.32
7663.25
10637.01
10921.02
11360.01
14041.27
15963.26

$\mu = \$2979.44$

$\sigma = \$7691.31$

Figure 5.3 Sample Observations from the Distribution
of Outcomes Associated with Strategy Two

v_1 = 160 acres of land rented

v_2 = 120 acres of corn to be planted

v_3 = 280 acres of soybeans to be planted

v_4 = May 26, the date after which all unplanted acreage is to
be planted in soybeans

The distribution of net cash income levels associated with this strategy is given in Figure 5.4. When compared to the distribution of income levels associated with the second strategy--the only member of the current efficient set--this distribution neither dominates nor is dominated. After three iterations, then, the efficient set is comprised of the second and third strategies.

The process continues in this manner until a prespecified number of iterations have been completed. As noted earlier, there is no guarantee that a true optimum will be identified. If a sufficient number of strategies are evaluated, however, it is almost certain that a nearly optimal strategy will be included in the efficient set. The number of iterations specified depends on the characteristic of the problem being analyzed. Donaldson and Webster (1968) suggest that 2000 strategies be examined. Experience to date with the relatively small problems considered in this study, however, indicates that 500 to 1000 iterations are often quite sufficient.

5.3.4 General Comments on the GREMP Model

The flexibility of the generalized procedure for the identification of preferred choices described above is its greatest strength. It can easily be adapted for use in the analysis of a diverse range of practical decision problems without requiring that major simplifying assumptions

-12468.33
-11453.37
- 8569.24
- 4272.46
- 740.21
4556.02
8818.28
9095.91
10008.77
10424.22
10543.07
10785.49
16644.11
18476.43
23776.90
25010.99
25050.58
26685.05
30877.08

$\mu = \$10,152.21$

$\sigma = \$12,516.70$

Figure 5.4 Sample Observations from the Distribution
of Outcomes Associated with Strategy Three

be made. With regard to problem formulation, the use of a random search procedure to generate strategies for consideration permits considerable flexibility with respect to both the types of control variables which can be specified in the definition of a management strategy and the types of constraints which can be imposed on them. Choice variables can be discrete or continuous. They can be controllable system input levels, feedback control rule parameters, or even indicators of the form of a feedback rule. Constraints on control variables can be linear or non-linear and can take forms more complex than those permitted in decision models based on mathematical programming.¹ With regard to the determination of system output distributions associated with alternative strategies, the use of Monte Carlo simulation techniques greatly facilitates the realistic representation of the stochastic processes by which the outcomes of particular choices are determined. In the simple example discussed above, price, yield, and time available for fieldwork are all considered explicitly as random factors which affect the outcome of any choice. Few restrictions are placed on the form of exogenous system input distributions, and the relationships among these factors, controllable system inputs, and system outputs can be quite complex in nature. With regard to the representation of decision maker preferences, the use of stochastic dominance with respect to a function also contributes greatly to the flexibility of the approach without sacrificing the logical power of decision theory based on the

¹A discussion of some of the types of constraints in control variables which are possible in the GREMP model is given in Appendix C.

expected utility hypothesis. The interval preference measurements used with this criterion can be as precise or imprecise as is required for any particular decision analysis.

The GREMP model is also relatively efficient computationally. In the analysis of one test problem with thirty-five choice variables and twelve linear constraints, for example, 1000 alternative strategies are generated and evaluated using less than seventy seconds of CPU time in a CDC6500. Furthermore, the core size of the computer program which implements the GREMP model is relatively small, and the degree of computational accuracy required for interval calculations is not unusually great. This suggests that it may be possible to design computer software which will permit the use of the GREMP model on a moderately sized personal computer.

Several criticisms of the GREMP model can be made. As was noted earlier, there is no guarantee that the model will identify a true optimum, since alternative strategies are generated in a random manner. Furthermore, the efficient set of choices may be quite large if the decision maker's preferences are not measured precisely enough. Neither of these criticisms is particularly serious, however. With regard to the first, in many instances it can be argued that a good solution to a well-formulated problem is preferable to an optimal solution to a problem which bears little resemblance to that actually facing the decision maker. The flexibility of the GREMP model, then, compensates for this weakness. With regard to the size of the efficient set, it can always be reduced by making more precise interval measurements of preferences, but it should be remembered that such reduction may lead to the exclusion of the decision maker's preferred choice from the efficient set.

A more valid criticism of the procedure is that it provides no direct information on how the composition of the efficient set might change if probability assessments were altered or if alternative values were assigned to certain key system design parameters. Such information can be obtained only by specifying the changes and repeating the procedure. For complex problems a sensitivity analysis conducted in this manner can be a costly and time-consuming process.

The GREMP model can also be criticized because it requires that the evaluation of alternative strategies be based on the distribution of a single system output variable, usually some measure of income or wealth output variable. In reality, most decision makers are concerned with more than one performance criterion. They have multiple objectives, and they consider trade-offs among these objectives when making choices. Decision theorists have focused considerable attention in recent years on the construction of preference measures which depend on more than one performance criterion and on the incorporation of such preference measures into a decision analysis. Unfortunately, however, the criterion of stochastic dominance with respect to a function has not been extended to the multivariate case, and there is no indication that such an extension will be made in the near future. Despite the acknowledged existence of multiple goals which affect choices, it can be argued that in many instances the consideration of only the most important objective may be adequate. The added accuracy attained from a more complete analysis may not justify the added cost of such an undertaking. When it is judged that more than one performance variable must be considered in a decision analysis, however, the GREMP model can be used if modified

slightly. All that is necessary is that a single-valued multiple criterion utility function be incorporated into the model to replace the evaluative component based on stochastic dominance with respect to a function.

Finally, two criticisms of a practical nature can be made of the model. First, it should be noted that, although the program which complements the GREMP model can be easily adapted for use in the analysis of a wide range of problems, the user is required to supply several problem-specific subroutines and so must have some programming skills. Though expertise in computer programming is certainly not required, this may preclude the use of the model in some instances. Second, it can be noted that in the analysis of complex decision problems with a large number of control variables, a large number of strategies must be examined before a nearly optimal one is identified. Computational costs can be considerable, then, for large problems. It should be noted, however, that with careful problem formulation and with the use of feedback control rules the size of the feasible set of strategies can be greatly reduced.

5.4 An Application

In this section the GREMP model is applied to the analysis of the sample problem discussed in the three preceding chapters. Efficient sets of choices are identified from a set of 500 feasible strategies for each of the three decision makers whose preference measures were given in Section 4.8 of Chapter IV. Since this sample problem was the basis for the discussion of the GREMP model in the preceding section, the reader should be familiar with its essential features. Therefore, the results of this application will be presented without further discussion.

The efficient set of choices for decision maker A is comprised of the eight strategies defined in Table 5.1.¹ Levels of land rented range from 0 to 160, with four of the eight strategies calling for land rental levels of 80 acres. Soybeans are the predominant crop in each plan, which is understandable given the cost-price relationships for the two crops. At low corn acreage levels, the feedback control rule parameter has little effect, so the switching date is of little importance in these strategies. Mean income levels realized under the efficient strategies range from slightly less than \$3000 to slightly above \$10,000. Minimum income levels vary little from one strategy to another. Maximum income levels, however, are significantly affected by land rental levels. It is also interesting to note that the efficient set need not be a mean-variance efficient set. For example, strategy 2 dominates strategy 4 by the mean-variance criterion, but both are in the decision maker's efficient set determined by stochastic dominance with respect to a function.

The efficient set of decision maker B is comprised of the nine strategies defined in Table 5.2. In this case, land rental levels tend to be higher than those called for by the strategies in decision maker A's efficient set. At higher acreage levels, the mix between corn and soybeans becomes more even, though most of the available acreage is planted to soybeans in each strategy. Mean net income levels tend to be higher within this set of strategies, but the dispersion of possible net income levels is also greater. Given the differences in the

¹System output levels are not enumerated for each state of nature. Rather, the mean, standard deviation, minimum value and maximum value are given for each distribution.

Table 5.1 Efficient Strategies for Decision Maker A

| Efficient Strategy | Control Variable Levels | | | | Properties of Net Cash Income Distribution | | | |
|--------------------|-------------------------|------------|----------------|----------------|--|--------------------|---------------|---------------|
| | Acres Rented | Acres Corn | Acres Soybeans | Switching Date | Mean | Standard Deviation | Minimum Value | Maximum Value |
| 1 | 0 | 0 | 240 | June 3 | 3816 | 8357 | -11875 | 20368 |
| 2 | 160 | 120 | 280 | May 26 | 10152 | 12517 | -12468 | 30877 |
| 3 | 80 | 60 | 260 | May 18 | 7138 | 10328 | -12102 | 25972 |
| 4 | 160 | 110 | 290 | May 26 | 9936 | 12526 | -12605 | 31170 |
| 5 | 80 | 50 | 270 | May 18 | 7168 | 10437 | -12220 | 26419 |
| 6 | 80 | 80 | 240 | June 3 | 6994 | 10142 | -11865 | 24305 |
| 7 | 0 | 50 | 190 | May 26 | 2949 | 7691 | -11482 | 15963 |
| 8 | 80 | 70 | 250 | June 3 | 7079 | 10239 | -11983 | 25165 |

Table 5.2 Efficient Strategies for Decision Maker B

| Efficient Strategy | Control Variable Levels | | | | Properties of Net Cash Income Distribution | | | |
|--------------------|-------------------------|------------|----------------|----------------|--|--------------------|---------------|---------------|
| | Acres Rented | Acres Corn | Acres Soybeans | Switching Date | Mean | Standard Deviation | Minimum Value | Maximum Value |
| 1 | 160 | 160 | 240 | May 26 | 9840 | 12199 | -12987 | 27685 |
| 2 | 240 | 200 | 280 | May 26 | 10808 | 15481 | -17272 | 33952 |
| 3 | 160 | 120 | 280 | May 26 | 10152 | 12517 | -12168 | 30877 |
| 4 | 160 | 140 | 260 | May 18 | 10167 | 12466 | -12231 | 30331 |
| 5 | 240 | 190 | 290 | May 26 | 10798 | 15483 | -17272 | 33952 |
| 6 | 160 | 130 | 270 | May 18 | 10193 | 12472 | -12350 | 30366 |
| 7 | 160 | 140 | 260 | June 3 | 10088 | 12395 | -12231 | 29610 |
| 8 | 160 | 130 | 270 | May 26 | 10167 | 12458 | -12350 | 30366 |
| 9 | 160 | 150 | 250 | June 3 | 9949 | 12367 | -12604 | 28764 |

preference measurements for decision makers A and B, the dissimilarity between these two efficient sets is understandable. The interval measurement of absolute risk aversion for decision maker A indicated a high level of risk aversion at negative net income levels--i.e. he has an apparently strong aversion to losses. Decision maker B, on the other hand, has much lower levels of absolute risk aversion at low income levels, and his efficient set contains strategies which, while providing opportunities for the realization of high income levels, also can result in substantial losses.

The efficient set of decision maker C is comprised of the strategies defined in Table 5.3. Of interest in this case is the fact that each of these eleven strategies is a member of the efficient set of either decision maker A or B. In a sense, then, it can be said that this decision maker's preference measurement lies between those of the other two decision makers.

Several general comments can be made about these results. First, they provide clear evidence of the discriminatory power of the preference measures based on the interval approach. The largest of three efficient sets contains only two percent of the total number of strategies examined. Second, the results demonstrate once again that preferences do have an important impact on the choices made by an individual. Finally, it should be observed that a search of only 100 feasible strategies identified many of the strategies included in these three efficient sets. This indicates that in some cases the evaluation of an extremely large number of strategies may not be necessary.

Table 5.3 Efficient Strategies for Decision Maker C

| Efficient Strategy | Control Variable Levels | | | Properties of Net Cash Income Distribution | | | | |
|--------------------|-------------------------|------------|----------------|--|-------|--------------------|---------------|---------------|
| | Acres Rented | Acres Corn | Acres Soybeans | Switching Date | Mean | Standard Deviation | Minimum Value | Maximum Value |
| 1 | 80 | 60 | 260 | May 18 | 7138 | 10328 | -12102 | 25792 |
| 2 | 160 | 140 | 260 | May 18 | 10167 | 12466 | -12231 | 30331 |
| 3 | 160 | 120 | 280 | May 26 | 10152 | 12517 | -12468 | 30877 |
| 4 | 160 | 110 | 290 | May 26 | 9936 | 12526 | -12605 | 31170 |
| 5 | 80 | 50 | 270 | May 18 | 7168 | 10437 | -12220 | 26419 |
| 6 | 80 | 80 | 240 | June 3 | 6994 | 10142 | -11865 | 24305 |
| 7 | 0 | 0 | 240 | June 3 | 3816 | 8357 | -11875 | 20368 |
| 8 | 80 | 70 | 250 | June 3 | 7079 | 10239 | -11983 | 25165 |
| 9 | 160 | 130 | 270 | May 18 | 10193 | 12472 | -12350 | 30366 |
| 10 | 160 | 140 | 260 | June 3 | 10088 | 12395 | -12231 | 29610 |
| 11 | 160 | 130 | 270 | May 26 | 10166 | 12458 | -12350 | 30366 |

CHAPTER VI
COMBINED PRODUCTION AND MARKETING DECISIONS BY
CASH GRAIN FARMERS: AN EXTENDED APPLICATION

6.1 Introduction

The formulation of an integrated set of operational techniques for the analysis of decisions made under uncertainty was the focus of the preceding chapters of this study. A simple example related to land rental and crop production decisions made by cash grain farmers has been used to illustrate these techniques. In this chapter the usefulness of the methodological tools developed above is demonstrated further by expanding the earlier example to include the consideration of alternative marketing strategies in conjunction with the selection of a cropping plan. Two modes of marketing will be evaluated: the sale of all production at harvest in the cash market and forward contracting.¹ The objectives here are to examine how these two modes of marketing can best be combined in the formulation of a marketing strategy which is appropriate for a particular decision maker, to determine the degree of interdependence between crop production and marketing strategies, and to examine the impact of changes in preferences on combined production and marketing strategies.

¹To simplify the discussion, other marketing alternatives such as hedging in the futures market and participation in government price stabilization programs will not be considered, nor will the use of crop storage as a marketing tool be examined. These alternatives could be incorporated into the analysis presented below, however.

Cash grain farmers make major resource allocation decisions under conditions characterized by uncertainty with respect to both prices and output levels. With the increased dependence of the feed grain sector on foreign markets in recent years, the impact of price uncertainty on cash grain producers has been particularly strong. Many have come to realize that marketing as well as production decisions have a major impact on the level of income they realize. The common practice of selling all production at harvest in the cash market often results in the receipt of prices which are low relative to those which could be realized under alternative modes of marketing. Furthermore, a strategy comprised only of cash sales at harvest does nothing to diminish the degree of price uncertainty faced by the producer over the period prior to and during planting--the period when major allocative decisions must be made. As a result, many producers find it desirable to consider forward pricing some or all of their planned production. By contracting to deliver a certain quantity of grain on a future date at a specified price, the producer sells all or some portion of his crop in advance. In doing so, he establishes with certainty a price for at least part of his total production, thereby greatly reducing the degree of price uncertainty he faces.

The reduction in price uncertainty achieved through forward contracting can be of considerable value in some situations. Advance knowledge of the price to be received simplifies the planning process. Furthermore, if the contract price is high enough, a producer who forward contracts may almost ensure that he will realize an adequate level of income. There are also costs associated with contracting, however.

The other party in the contract, the buyer of the grain, often has access to better information than that available to the producer and it is unlikely that he will offer a price higher than that he himself expects to receive. Of equal importance is the fact that the seller, though he protects himself against the effect of an unexpected downturn in price, also foregoes the opportunity to benefit from unexpected price increases. Finally, grain must be delivered according to the terms of the contract even if production falls short of the amount which is forward priced. If yields are unusually low or if poor weather conditions prevent the planting of some acreage altogether, the producer faces the prospect of being forced to purchase grain on the cash market to meet the terms of his contract. Given the advantages and disadvantages of forward contracting, then, the producer must determine how much, if any, of his anticipated production he wishes to contract. Clearly preferences, financial position, price expectations and production plans affect this decision.

The determination of optimal forward contracting levels for agricultural producers facing yield and price uncertainty has been examined in a mean-variance framework by McKinnan (1967). He shows that under relatively simple conditions the optimal level of forward contracting depends on five fundamental parameters: the standard deviations of crop yield and product price at harvest, the expected crop yield, the forward contract price (which is also the expected price at harvest), and the correlation coefficient between crop yield and harvest price. McKinnan's work has been extended by Ward and Fletcher (1971) and by Heifner (1972), and Barry and Willmann (1976) provide an interesting

empirical application of contracting theory based on mean-variance analysis. The applicability of the results of each of these studies is limited by the somewhat restrictive assumptions associated with the use of the mean-variance criterion--assumptions of normally distributed net returns and risk averse behavior. The requirement that net returns be normally distributed may be particularly unrealistic. More critical, however, is the failure in each study to consider a factor which is of primary importance in an applied context. All treat forward contracting decisions as though they can be made at only one point in time. In reality producers have forward pricing opportunities open to them over an extended period of time. In such a context the decision of when to contract may be as important as the decision of how much to contract. Producers continually evaluate forward pricing opportunities, and it is not unusual for an individual to enter into contracts at several different times. Once a contract is made, however, it must be honored. Choices made in the present, then, affect opportunities in the future. Therefore, it is important to consider forward contracting decisions in a more dynamic analytical framework.

The techniques developed in this study allow the relaxation of the restrictions on probabilities and preferences imposed by the use of mean-variance analysis. More important, however, they permit the evaluation of both production and forward contracting decisions in a more dynamic framework in which the roles of learning and adaptive behavior are treated more explicitly than in previous studies. In the analysis below, attention will again be focused on decisions made by a cash grain farmer producing corn and soybeans. As in the example discussed

in earlier chapters, flexibility is introduced into the production planning process through the incorporation of a simple stopping rule for corn planting. Marketing strategies for both corn and soybeans are defined by more complex feedback control rules which are applied repeatedly over a seven month period extending from mid-January to mid-August. Particular attention will be given to the examination of both the interdependence between production and marketing strategies and the impact of changes in decision maker preferences on the choice of a combined production-marketing strategy.

6.2 Problem Formulation

The basic decision situation in this extended example is the same as that described in Section 2.5 of Chapter II. The operator of a relatively small southeastern Michigan cash grain farm needs to realize a substantial level of income from his farming operation in order to meet his debt repayment commitments of \$35,000 annually and to cover family living expenses. He wishes to choose a management strategy which, given his risk preferences and the range of opportunities open to him, will best satisfy this need. The time is January 1979, and, because land rental decisions must be made, the farmer must formulate at least a tentative management strategy now.

The system of concern in this example is comprised of a set of production and marketing processes which constitute a farming operation. The performance of this system is measured by a single system output variable: annual net income for family living expenses and firm expansion after all debt repayment commitments have been met. The level of income realized is affected by the structure of the system and by exogenous and controllable system inputs.

The structure of the production and cash marketing processes embodied in this system was described in Section 2.5 and remains unchanged in this example. It will be recalled that standard crop budgets for corn and soybeans were presented and that relevant planting and harvest periods were specified for each crop, as were rules of thumb which determine the priority of each crop during planting and harvesting. Important state variables related to the production process included production expenses incurred to date, acreage of each crop planted or harvested to date, and bushels of each crop harvested to date. All production was sold in the cash market at harvest in the original example. The harvest price of each crop was multiplied by the number of bushels harvested to determine the value of marketing receipts for each crop, the key state variable used to describe the marketing process.

The incorporation of forward contracting into the analysis requires that several structural features associated with this mode of marketing be specified and that several new state variables be defined. Structurally, it must be recognized that contracts, once made, are binding, and that if production falls short of the amount contracted, enough grain must be purchased in the cash market at harvest to cover the deficiency. It must also be noted that, unlike a hedge in the futures market, a hedge based on a forward contract cannot be lifted. The new state variables which must be defined include the number of bushels of each crop contracted to date and the current level of receipts forthcoming at harvest from quantities of each crop contracted.¹

¹Several additional state variables will be defined during the discussion of the feedback control rules which determine the marketing strategies for corn and soybeans.

Exogenous system inputs in the original example included the following stochastic environmental factors: the price at harvest of each crop, the number of days available for fieldwork in any particular planting or harvest period, and the yield of each crop for each allowable planting-harvest combination. Consideration of forward contracting requires the specification of an additional set of stochastic factors: the harvest delivery forward contract prices for corn and soybeans over the period when contracting decisions are to be made. Clearly this set of prices will affect the desirability of forward contracting and the level of income realized under any management strategy which calls for the forward contracting of either crop, and clearly these prices cannot be known with certainty in mid-January when a management strategy must be formulated.

Only four control variables were considered in the example discussed in previous chapters. They were:

- v_1 = acres of land rented
- v_2 = acres of corn to be planted
- v_3 = acres of soybeans to be planted
- v_4 = the date after which all unplanted acreage is to be
planted in soybeans

The consideration of forward contracting decisions requires the specification of several new control variables, which will determine the number of bushels of each crop contracted at any particular time. Contracting levels, per se, will not be selected directly. Rather, a feedback control rule which determines the desired contracting level at any point in time will be specified for each crop. The form of this rule is the

same for both corn and soybeans and is similar to that discussed in Section 2.3.1 of Chapter II. It is:

$$DBC_t = EC_t(v_w ZA_t + v_x ZB_t + v_y ZC_t) \quad 6.1$$

where

DBC_t = desired bushels contracted at time t

EC_t = current expected size of harvest

ZA_t = the percentage difference between the current contract price, CP_t , and the current expected harvest price, EP_t ; i.e. $ZA_t = (CP_t - EP_t)/EP_t$

ZB_t = the current daily rate of change in the contract price; i.e. $ZB_t = dCP_t/dt$

ZC_t = the difference between the desired percentage of the expected crop contracted and the actual percentage contracted. The desired percentage contracted is defined by the expression $AP_t/DAP_t + v_z$, where AP_t and DAP_t are current actual and desired acreage planted and v_z is a parameter to be selected. The actual percentage contracted is defined by the expression BC_t/EC_t , where BC_t is the number of bushels contracted to date and EC_t is the expected size of the harvest. It follows that $ZC_t = AP_t/DAP_t + v_z - BC_t/EC_t$.

In addition to v_z , v_w , v_x , and v_y are parameters to be selected.

The inclusion of each term in this contracting rule can be justified by appealing to commonly recognized rules of thumb regarding forward contracting. The first term ZA_t , reflects an assessment of the fundamental position of the market. If ZA_t is positive, the current contract

price is above the current expected harvest price. As ZA_t becomes larger, the attractiveness of current pricing opportunities increases and the decision maker is expected to want to contract more. Similarly, if ZA_t is negative, the current pricing opportunity is not an attractive one and the desired level of contracting will be less. The parameter which weights this factor, v_w , is expected to be positive.

The second term, ZB_t , reflects, in a very simple way, a technical assessment of market conditions. If ZB_t is positive, the contract price is increasing, and most technical analysts would recommend that contracting be delayed. If ZB_t is negative, and if fundamental analysis indicates that the current pricing opportunity is a favorable one, on the other hand, many technical analysts would recommend that forward contracts be entered into. Following this reasoning, the parameter which weights this factor, v_x , is expected to be negative.

The third term, ZC_t , reflects the degree to which current contracting levels coincide with desired contracting levels. In this case desired levels are a linear function of the percentage of desired acreage actually planted. The parameter v_z shifts the intercept of that function and is expected to be negative, reflecting the fact that many decision makers hesitate to contract before planting begins unless pricing opportunities are particularly favorable. When ZC_t is positive, the decision maker is expected to desire to contract more unless other factors indicate the current pricing opportunity is not a good one. When ZC_t is negative, the desire to contract is diminished. Therefore, the parameter which weights the importance of this factor in the contracting rule, v_y , is expected to be positive.

As stated above, a control rule of this form is specified for each crop. Since four parameters must be selected for each rule, a total of eight new control variables are required in this example. Control variables v_5 , v_6 , v_7 , and v_8 correspond to v_w , v_x , v_y , and v_z for the corn contracting rule, and v_9 , v_{10} , v_{11} , and v_{12} correspond to the same parameters for the soybean contracting rule. Introduction of the contracting rules into the analysis also requires the specification of several new state variables. In addition to those mentioned earlier, state variables indicating the daily rate of change in the contract price, the expected size of the harvest, and the expected price at harvest must be monitored for each crop. The daily rate of change in the contract price is defined by the expression:

$$\frac{dCP_t}{dt} = (CP_t - CP_{t-1})/dt \quad 6.2$$

where CP_t and CP_{t-1} are successive observations of the contract price and dt is the number of days between price observations. The expected size of harvest for a particular crop is defined by the expression:

$$EC_t = (BY)(DAP_t) \quad 6.3$$

where DAP_t is the desired acreage to be planted and BY is a base yield.¹ The expected harvest price for a particular crop is defined by the following simple expectations model:

$$EP_t = .5CP_t + .33CP_{t-1} + .17CP_{t-2} \quad 6.4$$

This is simply a weighted moving average of the three most recent contract price observations. Finally, it should be noted that in the

¹ BY is set at 100 bushels per acre for corn and 33 bushels per acre for soybeans.

application of each rule, the number of bushels to be contracted at any time is set at zero if the number of bushels already contracted is equal to 150 percent of the expected crop.

As in the previous example, the planning and decision horizons are both said to be a single crop year. It will be recalled that the stopping rule for corn planting is checked at the end of the second, third, and fourth planting periods. The two contracting rules are consulted a total of twelve times--four times prior to the commencement of planting, at the end of each of the six planting periods, and twice between planting and harvest. The exact dates for each application of the two rules are given in Figure 6.1.

The problem in this example, then, is to identify a management strategy defined by a total of twelve control variables which maximizes the expected utility of a particular decision maker. Three of these choice variables specify controllable system input levels directly. The other nine are parameters in feedback control rules which are consulted periodically over the entire decision horizon.

6.3 The Determination of Subjective Probability Distributions

Subjective probability distributions for crop yields, time available for fieldwork, and crop prices at harvest were specified in Section 3.4 of Chapter III. They remain unchanged in this example. It is also necessary to specify subjective distributions for the harvest delivery contract price of each crop at each of the eleven dates after January 10 listed in Figure 6.1. These prices cannot be known with certainty at the time when a strategy must be selected, and their distributions can have an important impact on the distribution of net income levels realized under any particular management strategy being considered.

January 10
February 10
March 20
April 20
May 10
May 18
May 26
June 3
June 11
June 19
July 20
August 20

Figure 6.1 Dates for the Application of Forward Contracting Rules

For either crop, elements in the series of contract prices are expected to be correlated with each other and with the cash price at harvest. Therefore, individual contract price distributions cannot be specified independently. Even if this could be done, it is unlikely that a decision maker would have clearly formulated expectations concerning contract price levels at each of eleven dates over a seven month period. To simplify the specification of these distributions, then, the following procedure which reflects the author's own subjective assessments was used. The harvest delivery contract price for each crop on January 10, 1979, the date when a strategy is to be selected, is known with certainty. That for corn is \$2.08, and that for soybeans is \$6.31. For any particular state of nature, the cash price at harvest can also be specified, this price representing a sample observation from the underlying subjective price distribution. In addition, contract margins charged by elevators are relatively constant and can be set at \$.10 per bushel for corn and \$.25 per bushel for soybeans.¹ Given this information, it is assumed that the harvest delivery contract price offered on any date between January 10 and the contract delivery date can be adequately forecasted by linearly interpolating between the January 10 price and the price at harvest less the contracting margin.² In Figure 6.2, for example, the contract price of corn on January 10 is \$2.08. The cash price 295 days later November 1 is \$2.32. With a

¹Contract margins cover the elevator's operating expenses and represent a premium paid to the elevator by the farmer for the reduction in price uncertainty.

²Approximate delivery dates are October 15 for soybeans and November 1 for corn.

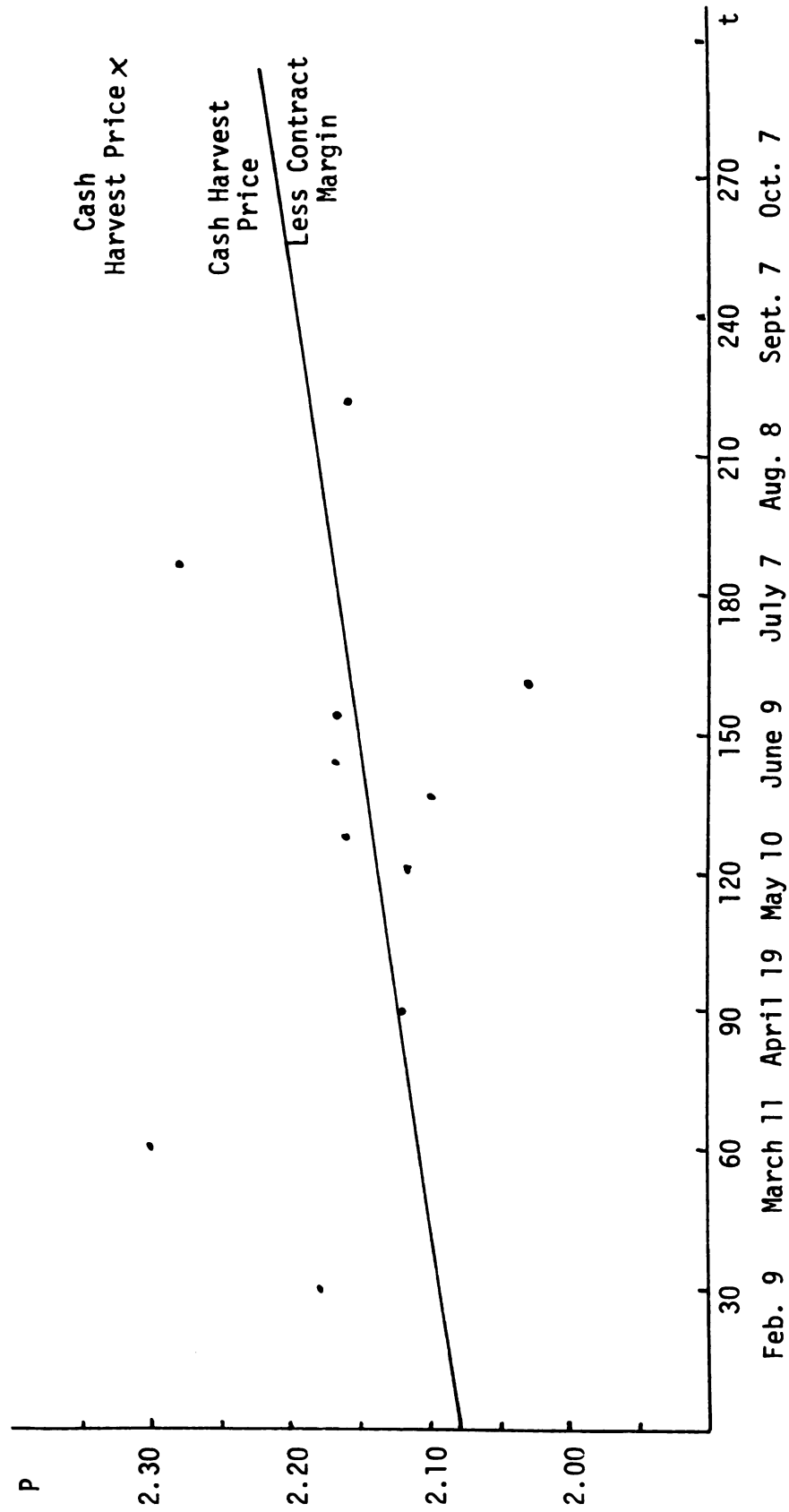


Figure 6.2 Price Forecast Line and Observed Control Price Levels for One State of Nature

contract margin of \$.10 per bushel, then, a hypothetical contract price offered on that day would be \$2.22. The line between the contract prices of \$2.08 on January 10 and \$2.22 on November 1 is used to forecast the contract price on intermediate dates. On May 10, 120 days after January 10, for example, the forecasted contract price is \$2.14. Forecasts of the contract price of each crop at each of the dates specified in Figure 6.1 can be made in this manner. By making such forecasts for a number of sample harvest price observations, a crude multivariate distribution of intermediate forward contract price levels can be specified.

Clearly forecasts based on such a simple model are subject to considerable error. To reflect this fact during the actual specification of sample observations from the contract price distributions, the forecast of each intermediate contract price was multiplied by a factor of $1+e$, where e is a normally distributed random variable with mean 0.0 and standard deviation .05. To further enhance the realism of the model, the multiplicative error terms for successive dates were correlated to reflect the fact that observed contract price levels are autocorrelated. A plot of values based on this more complex model is also shown in Figure 6.2. In this instance, relatively good pricing opportunities occur in March and July, and contract prices during the period when planting takes place cluster around the price forecast line. It is interesting to note that the contract price in this example never exceeds the actual cash price at harvest.

The distribution of net income levels realized under any particular management strategy is determined by simulating system performance under

that strategy in each of a set of sample states of nature.¹ One possible management strategy is defined by the following control variable levels: $v_1 = 240$ acres rented; $v_2 = 190$ acres corn; $v_3 = 290$ acres soybeans; $v_4 = \text{May 26}$, the date after which all unplanted acreage is planted in soybeans; $v_5 = 12$; $v_6 = 1$; $v_7 = .90$ and $v_8 = -.30$ for the corn contracting rule; and $v_9 = 17$, $v_{10} = -1$, $v_{11} = .50$, and $v_{12} = -.20$ for the soybean contracting rule. The sample observations which define the distribution of net cash income levels associated with this strategy are given in Table 6.1 along with the other information pertaining to system performance in each of the twenty states of nature considered. The cropping plan in this case is identical to that which maximized expected net returns in the earlier example which precluded forward contracting, and the net income levels realized under the strategy without contracting are also given in Table 6.1.

Several observations can be made about the differences in net income levels realized under these two strategies. First, the average net income level is higher under the strategy which precludes contracting. This is to be expected, given the contract margins on corn and soybeans of \$.10 and \$.25 respectively. Second, the variability of net income, as measured by the standard deviation of the sample observations, is reduced by forward contracting, again as would be expected. The reduction is not a sizeable one, however. It is also interesting to note that in state 14 when yields for both corn and soybeans are apparently

¹The computer program which implements the simulation model used in this phase of the analysis is subroutine DISGEN in the listing of the GREMP model at the end of Appendix C.

Table 6.1 System Performance under a Sample Management Strategy

| Net Income Levels with Contracting | Corn | | | | Soybeans | | | | Net Income Levels with No Contracting |
|------------------------------------|-------------------|--------------------|------------------------|--------------------|-------------------|--------------------|------------------------|--------------------|---------------------------------------|
| | Bushels Harvested | Bushels Contracted | Average Contract Price | Cash Harvest Price | Bushels Harvested | Bushels Contracted | Average Contract Price | Cash Harvest Price | |
| 11111.00 | 20071.01 | 4670.49 | 2.06 | 2.07 | 9554.04 | 1445.29 | 6.24 | 6.17 | 20849.20 |
| 12222.07 | 17140.77 | 5075.31 | 2.07 | 2.19 | 9105.27 | 1972.50 | 5.89 | 5.87 | 13210.50 |
| -6155.04 | 10210.53 | 10874.20 | 2.22 | 2.14 | 7078.92 | 5927.74 | 6.46 | 6.20 | -7572.00 |
| 20140.02 | 19212.64 | 2305.03 | 2.11 | 2.07 | 8713.70 | 461.27 | 6.31 | 6.69 | 14943.22 |
| 1454.28 | 9541.95 | 6586.00 | 2.10 | 2.30 | 8816.97 | 7523.00 | 6.34 | 6.19 | 1044.65 |
| 13150.38 | 16765.27 | 14730.87 | 2.18 | 2.13 | 8524.09 | 5376.50 | 6.76 | 6.29 | 2615.43 |
| 9922.35 | 14282.00 | 1508.04 | 2.11 | 2.07 | 9118.42 | 1560.07 | 6.12 | 6.20 | 10057.10 |
| 14385.86 | 17349.74 | 6116.33 | 2.24 | 2.22 | 9240.19 | 6587.10 | 6.40 | 6.26 | 12710.19 |
| -10614.82 | 15390.08 | 9020.90 | 1.99 | 1.87 | 7634.18 | 1801.64 | 5.57 | 4.55 | -13401.06 |
| 30479.53 | 18433.92 | 9604.11 | 2.01 | 2.03 | 9400.82 | 5117.10 | 7.16 | 7.51 | 33951.82 |
| 26451.02 | 18076.74 | 6127.94 | 2.22 | 2.31 | 8510.37 | 4668.36 | 7.17 | 7.63 | 24100.73 |
| 28571.65 | 17477.06 | 6852.73 | 2.31 | 2.31 | 9077.98 | 7681.69 | 7.20 | 7.35 | 24707.81 |
| -16582.61 | 7124.66 | 1330.30 | 1.94 | 1.83 | 8569.97 | 461.27 | 6.31 | 5.09 | -17271.70 |
| -17017.79 | 2909.32 | 2380.03 | 2.30 | 2.39 | 5738.02 | 5045.91 | 6.82 | 6.84 | -14474.02 |
| 4329.22 | 14002.06 | 4112.03 | 2.07 | 2.01 | 7174.74 | 3842.50 | 6.83 | 7.77 | 7705.34 |
| 6555.76 | 15324.04 | 6650.21 | 2.07 | 2.20 | 7838.63 | 2633.58 | 6.74 | 6.07 | 5656.63 |
| -331.42 | 16005.24 | 3577.33 | 2.11 | 1.91 | 8075.55 | 461.27 | 6.31 | 5.61 | -1344.75 |
| 4640.99 | 19212.60 | 9724.51 | 2.22 | 2.22 | 8784.70 | 461.27 | 6.31 | 5.54 | 13676.58 |
| 27879.44 | 20163.65 | 14310.72 | 2.21 | 2.20 | 9112.15 | 6312.83 | 6.50 | 7.05 | 30823.20 |
| 25305.76 | 16040.53 | 9464.51 | 2.26 | 2.32 | 9225.30 | 8183.35 | 7.07 | 7.15 | 26574.39 |

$\mu = \$10,868.57$

$\sigma = 15,381.73$

$\mu = \$10,797.51$

$\sigma = 15,482.77$

quite low, contracting costs the producer a considerable amount of money and turns an already bad situation into a worse one. The problem is particularly serious for corn. Pricing opportunities apparently appear to be quite favorable, since more corn is contracted than would normally be produced. At harvest, however, the cash price is \$.09 above the average price for which corn was forward contracted, and a total of 14,631 bushels must be purchased on the cash market at this higher price to meet the terms of the contracts. Finally, it should be noted that quantities of corn and soybeans contracted vary considerably from one state of nature to another. This reflects the fact that there are considerable differences in the attractiveness of the forward contracting opportunities available in each state. Clearly the contracting rule is sensitive to these differences.

6.4 Decision Maker Preferences

Interval preference measures for the three decision makers discussed in the illustrations presented in Chapters IV and V will be used to order the alternative strategies considered in this example. Orderings will also be made for a risk neutral decision maker--a decision maker whose absolute risk aversion function is always equal to zero--to determine whether forward contracting rules can be identified which lead to higher expected net income levels than those realized under strategies which preclude forward contracting.

The net income level distributions associated with the two strategies identified in the preceding section--one with contracting and one without--were ordered for each of the four decision makers considered in this example. The strategy which precludes forward contracting is

preferred to that with contracting by the risk neutral decision maker and by decision maker A. The two strategies cannot be ordered for decision makers B and C. It is interesting to note that the non-contracting strategy is preferred by decision maker A, whose preferences were most conservative in the previous example. This may be attributable to the aversion of this decision maker to the sizeable losses incurred under the contracting strategy in state 14 when crop yields are unusually low and contracting levels are high.

6.5 The Identification of Preferred Choices

By incorporating the simulation model described in Section 6.2 into the Monte Carlo risk programming model described in Chapter V, an efficient set of strategies can be identified for each of the four decision makers considered in this example. Each strategy specifies a flexible production plan and feedback control rules which direct forward contracting decisions for both corn and soybeans.

Constraints placed on the values of crop production and land rental control variables were discussed in Section 5.3 of Chapter V. They remain unchanged in this example. Only upper and lower bound constraints are placed on the parameters of the two contracting rules. It will be recalled from Section 6.2 that each rule has four parameters. Experimentation with the model indicates that the following ranges of admissible values are reasonable and do not unduly constrain the set of feasible choices:

$$\begin{array}{lll}
 0 \leq v_w \leq 20 & w=5, 9 & 6.5 \\
 -20 \leq v_x \leq 0 & x=6, 10 & \\
 0 \leq v_y \leq 1.0 & y=7, 11 & \\
 -.7 \leq v_z \leq 0 & z=8, 12 &
 \end{array}$$

Parameters v_w and v_x are restricted to integer values. Parameters v_y and v_z are restricted to values which are integer multiples of 0.1.

In this example 1000 randomly selected strategies were examined and ordered for each decision maker's preferences. This is twice the number of strategies considered in the earlier example. The fact that each strategy is now defined by twelve control variables rather than only four implies that many more feasible strategies exist, however, and therefore more strategies should be evaluated. The efficient sets identified range in size from a single element for the risk neutral decision maker and three elements for decision maker A to thirteen and seventeen elements for decision makers B and C respectively. It is not possible to discuss all the efficient strategies for each decision maker, therefore, a representative strategy from the efficient set of each decision maker is defined in Table 6.2.

The representative strategy for the risk neutral decision maker is that which maximizes expected net income under contracting. Detailed information on system performance under this strategy in each of the twenty states of nature considered is given in Table 6.3. The expected net return is slightly less than that realized under the expected net return maximizing strategy from the earlier example which precludes contracting, but in many respects the two strategies are remarkably similar. The two cropping plans are nearly identical,

Table 6.2 Representative Strategies from the Efficient Sets of Four Decision Makers

| Decision Maker | Acres Rented V_1 | Corn Acreage V_2 | Soybean Acreage V_3 | Switching Date V_4 | Corn Contract Rule Parameters | | | Soybean Contract Rule Parameters | | | | |
|----------------|-----------------------|-----------------------|--------------------------|-------------------------|-------------------------------|-------|-------|----------------------------------|-------|----------|----------|----------|
| | | | | | V_5 | V_6 | V_7 | V_8 | V_9 | V_{10} | V_{11} | V_{12} |
| Risk Neutral | 240 | 180 | 300 | May 26 | 20.0 | -16.0 | .20 | - .30 | 17.0 | -17.0 | .10 | - .20 |
| A | 0 | 0 | 240 | June 3 | ----- | ----- | ----- | ----- | 17.0 | - 2.0 | .80 | 0 |
| B | 160 | 140 | 260 | May 18 | 19.0 | - 5.0 | .60 | - .20 | 6.0 | - 2.0 | 0 | - .40 |
| C | 80 | 50 | 270 | May 26 | 16.0 | -19.0 | .90 | - .30 | 16.0 | -14.0 | .20 | - .60 |

Table 6.3 System Performance under the Preferred Strategy of
a Risk Neutral Decision Maker

| Net Income Levels with Contracting | Corn | | | | Soybeans | | | | Net Income Levels with No Contracting |
|--|----------------------|-----------------------|------------------------------|--------------------------|----------------------|-----------------------|------------------------------|--------------------------|---|
| | Bushels Harvested | Bushels Contracted | Average Contract Price | Cash Harvest Price | Bushels Harvested | Bushels Contracted | Average Contract Price | Cash Harvest Price | |
| 22611.17 | 21031.21 | 659.13 | 2.02 | 2.07 | 9554.04 | 0.00 | 1 ^a | 6.17 | 22611.17 |
| 13151.23 | 17467.99 | 659.13 | 2.02 | 2.19 | 9290.47 | 242.19 | 6.05 | 5.67 | 13151.23 |
| 10245.34 | 10210.53 | 5044.25 | 2.20 | 2.44 | 7079.92 | 0.00 | 1 | 6.20 | 10245.34 |
| 20010.42 | 18214.42 | 659.13 | 2.06 | 2.07 | 8948.35 | 0.00 | 1 | 6.48 | 20010.42 |
| 2495.12 | 9541.05 | 659.13 | 2.04 | 2.30 | 8816.87 | 2315.06 | 6.62 | 6.15 | 1744.65 |
| 13380.77 | 15927.29 | 6619.52 | 2.22 | 2.13 | 8773.79 | 2630.68 | 6.36 | 6.39 | 12686.61 |
| 10047.62 | 14282.09 | 659.13 | 2.08 | 2.07 | 9118.42 | 0.00 | 1 | 6.20 | 10057.10 |
| 13285.28 | 16912.52 | 659.13 | 2.08 | 2.22 | 8439.28 | 4178.96 | 6.50 | 6.26 | 12744.61 |
| 13460.49 | 14599.39 | 659.13 | 2.08 | 1.87 | 7815.20 | 0.00 | 1 | 4.55 | 13612.91 |
| 34136.72 | 14095.60 | 659.13 | 2.09 | 2.03 | 9427.06 | 0.00 | 1 | 7.51 | 34166.76 |
| 28656.56 | 17160.74 | 659.13 | 2.08 | 2.31 | 8675.47 | 41.02 | 7.03 | 7.67 | 28832.77 |
| 29613.99 | 16633.44 | 659.13 | 2.08 | 2.31 | 9289.58 | 0.00 | 1 | 7.35 | 29765.50 |
| 17106.92 | 7124.66 | 659.13 | 2.08 | 1.83 | 8369.97 | 0.00 | 1 | 5.09 | 17271.70 |
| 15306.33 | 8506.50 | 7594.67 | 2.37 | 2.35 | 5888.50 | 2621.93 | 6.52 | 6.88 | 14227.45 |
| 7452.25 | 13560.28 | 659.13 | 2.08 | 2.01 | 7174.74 | 0.00 | 1 | 7.77 | 7406.12 |
| 5791.43 | 15274.04 | 5037.12 | 2.23 | 2.20 | 7838.63 | 0.00 | 1 | 6.07 | 5658.63 |
| 1260.94 | 15112.54 | 659.13 | 2.04 | 1.51 | 8284.05 | 0.00 | 1 | 5.61 | 1402.09 |
| 13185.26 | 18664.16 | 659.13 | 2.02 | 2.22 | 881.94 | 0.00 | 1 | 5.54 | 13249.14 |
| 30666.34 | 19140.26 | 659.13 | 2.08 | 2.20 | 9356.55 | 215.59 | 6.07 | 7.09 | 30965.34 |
| 25665.07 | 16040.53 | 659.13 | 2.08 | 2.32 | 9225.39 | 5417.99 | 7.05 | 7.15 | 26376.39 |

^aThe average contract price is undefined when nothing is contracted.

$\mu = \$10,756.27$

$\sigma = 15,425.27$

$\mu = \$10,720.89$

$\sigma = 15,475.31$

and the contracting rules employed in the strategy which permits forward pricing do not result in extensive forward contracting for either crop in most states of nature. The parameter weighting the first term of each contracting rule is assigned a relatively high value, which implies that contracting is strongly discouraged when the contract price is below the expected harvest price and strongly encouraged when it is above the expected harvest price. This is precisely the behavior one would expect from an individual who seeks to maximize expected net returns. The parameter weighting the second term of each contracting rule is assigned a strongly negative value. This implies that contract price movements also have an important impact on this decision maker's choices. Even if the pricing opportunity is a favorable one, he will delay his commitment to a forward contract if the contract price is rising. Again this conforms to intuitive expectations about the behavior of such a decision maker. Finally, the third parameter of each contracting rule is relatively small, which implies that this decision maker places little emphasis on achieving a desired level of contracting under almost any market conditions.

Net income levels realized under an identical production strategy which precludes contracting are also given in Table 6.3. Comparison of the two net income distributions indicates that contracting can lead to both higher expected net returns and less income variation, though actual differences between the two strategies are minimal. It is also interesting to note that the corn contracting rule is particularly successful in this instance, the average corn contracting price being well above the cash price at harvest in three of the four states of

nature when more than 1000 bushels are contracted. The soybean contracting rule is somewhat less successful.

Information on system performance under the preferred strategy of decision maker A is given in Table 6.4. This strategy is markedly different from that of the risk neutral decision maker. Its production component calls for no land rental and for the planting of all available acreage in soybeans. This is identical to one of the efficient production plans identified for this decision maker in the earlier example. Because no corn is to be planted, only the soybean contracting rule has any impact on system performance. As the information given in Table 6.4 indicates, this rule results in substantial levels of forward contracting in most states of nature. The parameter weighting the first term of the soybean contracting rule, v_9 , is relatively large, indicating that this decision maker also places considerable emphasis on fundamental analysis--on differences between the forward contract price and the expected cash price at harvest. Unlike the risk neutral decision maker, however, decision maker A assigns relatively little weight to technical analysis--to the analysis of contract price movements--as is indicated by the small absolute magnitude of parameter v_{10} . Also in contrast to the strategy of the risk neutral decision maker, the relatively large value of parameter v_{11} indicates that much stronger emphasis is placed on attaining some desired level of contracting under most conditions, and the fact that v_{12} is assigned a value of 0.0 implies that this desired level is 100 percent of the expected crop once all acreage has been planted.

Table 6.4 System Performance under the Representative Strategy for Decision Maker A

| Net Income Levels with Contracting | Corn | | | Soybeans | | | Net Income Levels with No Contracting |
|------------------------------------|-------------------|--------------------|------------------------|-------------------|--------------------|------------------------|---------------------------------------|
| | Bushels Harvested | Bushels Contracted | Average Contract Price | Bushels Harvested | Bushels Contracted | Average Contract Price | |
| 7173.05 | 0.00 | 0.00 | 1 ^a | 7864.30 | 4337.74 | 6.22 | 6951.10 |
| 4011.53 | 0.00 | 0.00 | 1 | 7866.49 | 4710.67 | 5.88 | 3014.65 |
| -3721.21 | 0.00 | 0.00 | 1 | 5589.04 | 7385.45 | 6.35 | -6754.79 |
| 6486.54 | 0.00 | 0.00 | 1 | 7456.06 | 3356.40 | 6.20 | 8098.64 |
| 4654.43 | 0.00 | 0.00 | 1 | 7405.29 | 8424.77 | 6.25 | 4139.00 |
| 4292.68 | 0.00 | 0.00 | 1 | 7242.43 | 7460.70 | 6.35 | 4601.17 |
| 2197.92 | 0.00 | 0.00 | 1 | 7129.53 | 3924.19 | 6.11 | 2540.96 |
| 5160.32 | 0.00 | 0.00 | 1 | 7139.52 | 9105.04 | 6.49 | 3029.85 |
| -7509.81 | 0.00 | 0.00 | 1 | 6528.10 | 4312.83 | 5.56 | -11875.09 |
| 14240.72 | 0.00 | 0.00 | 1 | 7454.79 | 7277.97 | 7.07 | 20377.68 |
| 10189.57 | 0.00 | 0.00 | 1 | 7257.27 | 6105.07 | 7.06 | 43692.98 |
| 13199.87 | 0.00 | 0.00 | 1 | 7745.67 | 9413.92 | 7.14 | 15182.27 |
| -5262.01 | 0.00 | 0.00 | 1 | 6633.52 | 1972.52 | 6.09 | -7828.05 |
| -1514.18 | 0.00 | 0.00 | 1 | 4900.14 | 7306.57 | 6.76 | -7637.09 |
| 2452.96 | 0.00 | 0.00 | 1 | 6263.91 | 4733.72 | 6.78 | 7129.61 |
| 221.36 | 0.00 | 0.00 | 1 | 6674.42 | 4312.64 | 6.37 | -1084.70 |
| -1483.91 | 0.00 | 0.00 | 1 | 6064.63 | 3209.91 | 5.95 | -2567.40 |
| 1832.54 | 0.00 | 0.00 | 1 | 7509.79 | 2894.06 | 6.05 | 365.47 |
| 1622.70 | 0.00 | 0.00 | 1 | 7731.99 | 8317.24 | 6.58 | 13073.31 |
| -0630.67 | 0.00 | 0.00 | 1 | 7500.04 | 9041.55 | 7.02 | 11911.57 |

^aThe average contract price is undefined when nothing is contracted.

$\mu = \$3328.75$

$\sigma = \$477.65$

$\mu = \$3815.71$

$\sigma = \$357.33$

Comparison of the net income levels realized under this strategy with those associated with a management strategy which precludes contracting but has an identical production plan is of considerable interest. The information given in Table 6.4 indicates that both the probability of negative net income levels and the magnitude of losses are reduced by the contracting rule. The probability of realizing relatively high income levels is also reduced, however, and expected net income is considerably higher under the strategy without contracting. When these two strategies are ordered using decision maker A's preference measurement, the strategy with contracting dominates that without.

Information on system performance under the representative strategy from the efficient set of decision maker B is given in Table 6.5.¹ In this instance 160 acres are rented, 140 acres are to be planted in corn, 260 acres are to be planted in soybeans, and May 18 is the date after which all unplanted acreage is planted to soybeans. The two contracting rules are quite different in this strategy. That for corn places a strong emphasis on differences between the current contracted price and the expected cash harvest price, the value of parameter v_5 being 19.0. Relatively little emphasis is placed on price movements, as evidenced by the small value of v_6 , and a rather strong emphasis is placed on contracting a specified percentage of the projected crop. This rule results in moderate to high contracting levels in thirteen states and is quite successful, with the average contract price being above the cash

¹The preferred strategy of the risk neutral decision maker is also in decision maker B's efficient set.

Table 6.5 System Performance under the Representative Strategy for Decision Maker B

| Net Income Levels with Contracting | Corn | | | | Soybeans | | | | Net Income Levels with No Contracting |
|------------------------------------|-------------------|--------------------|------------------------|--------------------|-------------------|--------------------|------------------------|--------------------|---------------------------------------|
| | Bushels Harvested | Bushels Contracted | Average Contract Price | Cash Harvest Price | Bushels Harvested | Bushels Contracted | Average Contract Price | Cash Harvest Price | |
| 10417.45 | 47427.00 | 1111.55 | 2.15 | 2.07 | 8416.56 | 236.81 | 6.34 | 6.17 | 1233.37 |
| 11617.11 | 13471.15 | 1029.27 | 2.13 | 2.19 | 9460.89 | 236.81 | 6.31 | 6.27 | 11517.31 |
| -3573.14 | 4619.06 | 5415.92 | 2.22 | 2.14 | 6334.75 | 236.81 | 6.31 | 6.20 | -4222.44 |
| 14706.07 | 14307.33 | 1022.22 | 2.18 | 2.07 | 7970.66 | 236.81 | 6.31 | 6.24 | 14576.85 |
| 11027.38 | 10639.80 | 2136.72 | 2.25 | 2.30 | 9161.30 | 750.72 | 6.56 | 6.10 | 10433.92 |
| 11014.67 | 12491.61 | 0293.63 | 2.23 | 2.13 | 7770.43 | 656.23 | 6.34 | 6.35 | 10105.07 |
| 3849.52 | 11349.79 | 0.00 | 1 ^a | 2.07 | 8199.88 | 236.81 | 6.31 | 6.20 | 2823.47 |
| 10780.80 | 12435.72 | 0.00 | 1 | 2.22 | 7337.56 | 1172.94 | 6.46 | 6.24 | 10544.26 |
| -11645.67 | 11408.00 | 992.12 | 2.04 | 1.87 | 6972.50 | 236.81 | 6.31 | 4.55 | -12231.11 |
| 30082.27 | 13514.77 | 2582.75 | 2.00 | 2.03 | 8717.50 | 273.75 | 6.44 | 7.91 | 30330.95 |
| 23954.62 | 13252.11 | 0.00 | 1 | 2.31 | 7802.47 | 236.81 | 6.31 | 7.63 | 24267.21 |
| 26237.59 | 12015.94 | 0.00 | 1 | 2.31 | 8717.87 | 1011.12 | 7.02 | 7.35 | 26574.02 |
| -8335.92 | 8413.04 | 0.00 | 1 | 1.83 | 7837.43 | 236.81 | 6.31 | 5.09 | -6624.83 |
| -12203.00 | 6711.57 | 17220.00 | 2.40 | 2.39 | 5243.54 | 347.31 | 6.43 | 6.92 | -12220.33 |
| 7432.03 | 10561.98 | 0.00 | 1 | 2.01 | 6859.51 | 236.81 | 6.31 | 7.77 | 7777.62 |
| 5739.29 | 11470.54 | 3735.45 | 2.25 | 2.20 | 7971.70 | 266.63 | 6.69 | 6.07 | 5017.30 |
| -737.11 | 12340.27 | 0.00 | 1 | 1.81 | 7439.63 | 236.81 | 6.31 | 5.61 | -522.67 |
| 11539.07 | 14576.53 | 3270.47 | 2.20 | 2.22 | 8101.19 | 236.81 | 6.31 | 5.54 | 11120.17 |
| 24551.40 | 14649.43 | 9124.77 | 2.24 | 2.20 | 8284.90 | 580.20 | 6.49 | 7.09 | 24545.50 |
| 23324.25 | 11624.27 | 4947.56 | 2.28 | 2.32 | 8633.00 | 1327.10 | 6.38 | 7.15 | 23761.18 |

^aThe average contract price is undefined when nothing is contracted.

$\mu = \$10,325.96$
 $\sigma = 12,277.87$

$\mu = \$10,167.49$
 $\sigma = 12,466.35$

harvest price in ten of the thirteen states in which there is contracting. The soybean contracting rule, on the other hand, is characterized by low values for each of its four parameters. It results in contracting levels below 1000 bushels in all but three states of nature and so has little effect on net income levels realized.

Comparison of the distribution of net income levels realized under this strategy with that associated with an identical production plan without contracting indicates that contracting results in a slightly higher expected net return and in some reduction in income variability. Given decision maker B's preferences, the strategy with contracting dominates that without.

Information on system performance under the representative strategy for decision maker C is given in Table 6.6.¹ In this case the production plan calls for the rental of 80 acres, for 50 acres of corn, and for 270 acres of soybeans. May 25 is the date after which all unplanted acreage is planted to soybeans, but the switching rule is inactive due to the low level of desired corn acreage. The contracting rules for both crops place considerable weight both on differences between the current contract price and the expected cash harvest price and on contract price movements. The corn contracting rule emphasizes the contracting of a certain percentage of the crop in most states of nature, however, while the soybean rule does not. Neither rule results in substantial contracting levels in most states of nature, though the

¹The representative strategy of decision maker B is also in decision maker C's efficient set.

Table 6.6 System Performance under the Representative Strategy for Decision Maker C

| Net Income Levels with Contracting | Corn | | | | Soybeans | | | | Net Income Levels with Contracting |
|------------------------------------|-------------------|--------------------|------------------------|--------------------|-------------------|--------------------|------------------------|--------------------|------------------------------------|
| | Bushels Harvested | Bushels Contracted | Average Contract Price | Cash Harvest Price | Bushels Harvested | Bushels Contracted | Average Contract Price | Cash Harvest Price | |
| 2545.56 | 6646.56 | 772.01 | 2.15 | 2.07 | 8462.76 | 0.00 | 1 ^a | 6.17 | -24.622 |
| 6715.25 | 4751.50 | 1055.38 | 2.14 | 2.19 | 9719.95 | 0.00 | 1 | 6.67 | 6965.71 |
| -5444.33 | 4171.50 | 3514.73 | 2.22 | 2.14 | 6104.08 | 0.00 | 1 | 6.20 | -5715.14 |
| 12744.60 | 5126.00 | 1756.08 | 2.18 | 2.07 | 9241.44 | 0.00 | 1 | 6.68 | 12551.52 |
| 2846.98 | 4876.56 | 1752.61 | 2.20 | 2.30 | 8657.10 | 1807.06 | 6.62 | 6.19 | 9216.06 |
| 8006.53 | 4457.50 | 4054.73 | 2.10 | 2.13 | 9024.60 | 2216.75 | 6.36 | 6.35 | 7030.65 |
| 5938.35 | 4940.00 | 242.10 | 2.06 | 2.07 | 7874.72 | 0.00 | 1 | 6.20 | 5041.27 |
| 7256.78 | 4507.50 | 946.42 | 2.14 | 2.22 | 7841.98 | 3525.90 | 6.50 | 6.26 | 6489.82 |
| -12077.71 | 4136.15 | 750.25 | 2.04 | 1.87 | 7154.70 | 0.00 | 1 | 4.55 | -12220.36 |
| 26448.10 | 5167.00 | 991.88 | 2.06 | 2.03 | 8732.65 | 0.00 | 1 | 7.91 | 26418.79 |
| 19524.24 | 4603.11 | 251.98 | 2.20 | 2.31 | 8075.07 | 0.00 | 1 | 7.63 | 19552.50 |
| 21167.63 | 4840.56 | 1126.27 | 2.25 | 2.31 | 8583.57 | 505.54 | 7.26 | 7.25 | 21270.03 |
| -7544.54 | 3932.30 | 258.97 | 1.95 | 1.83 | 7390.82 | 0.00 | 1 | 5.09 | -7275.62 |
| -9246.88 | 2434.15 | 2610.52 | 2.40 | 2.39 | 5415.24 | 1169.79 | 6.70 | 6.88 | -9055.42 |
| 8666.75 | 3771.64 | 263.33 | 2.05 | 2.01 | 6917.31 | 0.00 | 1 | 7.77 | 3676.22 |
| 1531.43 | 4493.00 | 1320.30 | 2.18 | 2.20 | 7311.65 | 0.00 | 1 | 6.07 | 1551.93 |
| -1028.93 | 4565.50 | 859.30 | 2.15 | 1.91 | 7654.55 | 0.00 | 1 | 5.61 | -1245.17 |
| 4854.50 | 5204.00 | 017.30 | 2.28 | 2.22 | 8344.64 | 0.00 | 1 | 5.54 | 4795.20 |
| 293.50 | 5133.50 | 1628.92 | 2.26 | 2.20 | 8561.49 | 0.00 | 1 | 7.05 | 4834.00 |
| 17760.45 | 4871.31 | 233.97 | 2.28 | 2.32 | 9340.08 | 3001.90 | 7.12 | 7.15 | 17975.56 |

^aThe average contract price is undefined when nothing is contracted.

$\mu = \$ 7,270.38$

$\sigma = 10,410.47$

$\mu = \$ 7,167.87$

$\sigma = 10,437.27$

corn rule does lend to contracting more than half the number of bushels harvested in three states of nature. The average contract price for corn exceeds the cash harvest price in twelve of the twenty states of nature, while that for soybeans exceeds the cash harvest price in only two of the six states in which there is contracting.

The distribution of net income levels realized under this strategy is quite similar to that associated with an identical cropping plan and no contracting. Expected returns are slightly higher under the strategy with contracting, however, and, given decision maker C's preferences, it is preferred to that without contracting.

6.6 Further Discussion of the Results

Several more general observations can be made regarding these results. First, they demonstrate once again that decision maker preferences have an important impact on the choice of a management strategy. Nowhere is this more evident than in the differences between the preferred strategy of the risk neutral decision maker and the representative strategy from the efficient set of decision maker A, whose level of absolute risk aversion is high over the negative range of net income values. Both cropping and marketing strategies are quite different, as are the associated net income distributions. The preferred strategy of the risk neutral decision maker calls for the rental of 240 acres and for a relatively balanced mix of corn and soybean production; it calls for the application of contracting rules which lead to an avoidance of forward pricing except in instances when pricing opportunities appear to be particularly favorable. The representative strategy for decision maker A, on the other hand, calls for no land rental and specialization

in soybean production and for the application of a contracting rule which results in the forward pricing of a substantial portion of the expected crop under nearly all market conditions. The expected net income level of \$10,756.27 under the preferred strategy of the risk neutral decision maker is more than \$7,400 higher than that associated with the representative strategy of decision maker A, but the latter is much less likely to face substantial losses under his preferred strategy.

The results also show that the introduction of forward contracting has only a minor impact on crop production strategies. Levels of land rental, relative acreages allotted to corn and soybeans, and stopping dates for corn planting specified in the strategies included in the efficient set of each decision maker show few significant changes. The most notable change is the inclusion of several cropping plans which devote more acreage to corn in the efficient set of decision maker C when forward contracting is incorporated into the management strategy.

A third observation is that, though many of the strategies identified appear to have some minor flaw, efforts to construct improved strategies through a search of a larger number of strategies or by simply changing parameter values which seem to cause problems rarely resulted in the identification of strategies with substantially better performance. In an effort to identify a better strategy for the risk neutral decision maker, for example, 250 additional strategies for which land rental is set at 240 acres were evaluated. The strategy defined in Figure 6.3 was identified as that which maximizes expected net returns. It has an expected net income level of \$10,868.57. This is higher than that of

Crop Production Plan

$v_1 = 240$ acres rented

$v_2 = 190$ acres corn

$v_3 = 290$ acres soybeans

$v_4 =$ May 26, the date after which
all unplanted acreage is
planted to soybeans

Corn Contracting Rule Parameters

$v_5 = 20.00$

$v_6 = -12.00$

$v_7 = .20$

$v_8^2 = -.40$

Soybean Contracting Rule Parameters

$v_9 = 7.00$

$v_{10} = -4.00$

$v_{11} = 0$

$v_{12}^2 = -.30$

Expected Net Return: \$10,868.57

Standard Deviation of Net Return: \$15,381.73

Lowest Sample Net Return: \$-17,108.70

Highest Sample Net Return: \$33,857.36

Figure 6.3 The Expected Net Return Maximizing Strategy

the preferred strategy identified in the original searches of strategies with and without contracting. The increase in the level of expected net income is not a very significant one, however.

Finally, it is interesting to note the composition of the efficient set for each decision maker when strategies with and without contracting are evaluated simultaneously. For the risk neutral decision maker the results based on the original search indicate that a strategy without contracting--that defined in Table 6.1--is preferred to all others. The strategy defined in Figure 6.3, which does involve contracting, has a still higher expected net return, however, and so is preferred by this decision maker.¹ For decision maker A, it was found that each of the eight efficient non-contracting strategies is dominated by at least one of the efficient strategies with contracting. For decision maker B seven of the nine efficient non-contracting strategies were dominated by at least one efficient contracting strategy, and for decision maker C all eleven efficient non-contracting strategies are dominated. In general, then, strategies with contracting clearly outperform those which preclude it under the conditions specified in this example.

6.7 Implications for Further Research

The problem of identifying combined production and marketing strategies which are well suited for a particular decision maker's situation is an important one. The analysis presented above is not intended to be a source of definitive solutions to this problem. Rather, it demonstrates an approach to the analysis of choices of this sort which employs

¹A more intensive search of non-contracting strategies failed to identify any with a higher expected net income level.

the techniques developed in this study to identify reasonable management strategies which perform well under relatively realistic conditions. More work needs to be done, however, to make the model described above the basis for truly reliable prescriptions.

With regard to problem formulation, other marketing alternatives need to be incorporated into the analysis. The impact on individual management strategies of government price stabilization programs which put an effective floor on crop prices and provide a form of disaster insurance for participants may be particularly important, for example. Similarly, the use of the futures market rather than forward contracting may be an attractive alternative for larger producers. In addition to the inclusion of other marketing alternatives, more careful specification of the feedback control rules which direct marketing decisions may also make the model a more reliable prescriptive tool.

Improved assessments of probability distributions for all the stochastic factors considered in this example are also needed. Particular attention should be given to the specification of more realistic distributions for forward contract price offerings over the course of the planning horizon. As was done in the example above, this can, perhaps, be best achieved by modelling the relationships among current contract price, cash price at harvest, and intermediate contract price offerings. In addition to improved specifications of subjective probability distributions, it may also be desirable to consider more than twenty sample states of nature in a decision situation such as this one in which a large number of random factors have an impact on net income levels.

Once modifications of this sort have been made, a more systematic exploration of the effect of decision maker preferences, probability assessments, and scale of operation on the nature of preferred forward pricing rules would be of particular interest. Can shifts in the parameter values of a specific rule be related to differences in factors such as these? Are contracting rule parameters relatively insensitive to changes in some of these factors? Is the choice of an expectations model dependent upon preferences? These are important questions which should be considered in future research. The objective of such research should not necessarily be to identify invariant behavioral rules which can be applied in any situation. The results above indicate that such rules are not likely to exist. Rather, the goal should be to better understand both the contracting rules themselves and the complex interactions among the variety of factors which affect decision maker choices.

CHAPTER VII

SUMMARY AND CONCLUSIONS

7.1 A Review of the Methodological Tools Developed in this Study

This study has focused on the development of techniques designed for use during four important phases of an applied decision analysis: problem formulation, the determination of subjective probability distributions, the measurement of decision maker preferences, and the identification of preferred choices. When considered together, the procedures described in this study represent an integrated set of operational techniques which facilitate the application of powerful theoretical tools based on the expected utility hypothesis in an applied setting.

In the discussion of problem formulation, two important considerations were stressed: the need to structure the problem being analyzed by identifying and classifying factors which will have an important impact on the outcome of the decision being made, and the need to give careful attention to the definition of what is to be decided. With regard to the first of these considerations, the usefulness of systems identification (Manetsch and Park, 1977a) as an aid in structuring the decision maker's perception of a particular practical problem was demonstrated. With regard to the definition of what is to be decided, the need to recognize the existence of future opportunities for learning in many decision situations and the desirability of flexible management

strategies in such instances were emphasized. The incorporation of feedback control rules into a management strategy was shown to be one way in which considerations of flexibility can be introduced into a decision analysis.

The value of combining direct probability assessments of underlying stochastic factors with the modelling of more complex stochastic processes was stressed in the discussion of the determination of the distribution of outcomes associated with any particular choice. Under this approach, which is suggested by Spetzler and Stael von Holstein (1975), direct encoding techniques are used to elicit information on a decision maker's expectations about future levels of environmental variables which cannot be controlled by the decision maker but have an important impact on the outcome realized under any particular strategy. Monte Carlo simulation techniques are then used to determine the combined effect of these exogenous stochastic factors and a particular management strategy in the distribution of system output levels. The use of both direct encoding and modelling allows considerable flexibility in the specification of subjective probability distributions of underlying variables and in the representation of the stochastic system under consideration--flexibility that is often lost when a more strictly analytical approach is taken.

One important criticism of systems modelling and Monte Carlo simulation is that correlations between random factors are often ignored. This is due largely to the fact that procedures for the generation of sample observations from multivariate distributions have been developed for only a few special distributions. An important

contribution of this study is the development of a generalized multivariate process generator, which is described in detail in Appendix A. This greatly enhances the value of Monte Carlo simulation as a tool in the applied analysis of decisions made under uncertainty.

The need for preference measurement techniques which are more flexible and more reliable than existing procedures is stressed in the discussion of the measurement of decision maker preferences. In response to this need, a new approach to preference measurement has been developed as a part of this study. This new procedure permits the construction of interval measurements of a decision maker's absolute risk aversion function. Perhaps its most important feature is that it allows direct specification of the degree of precision with which preferences are to be measured. In contrast, single valued utility functions are exact but often inaccurate preference measures, and commonly used stochastic efficiency criteria are based on inexact, inflexible assumptions about preferences.

Interval measurements of absolute risk aversion are used in conjunction with the criterion of stochastic dominance with respect to a function (Meyer, 1977a) to order alternative choices. This recently developed stochastic efficiency criterion can be used to evaluate alternative choices for classes of decision makers defined by upper and lower bound absolute risk aversion functions. As such, it is both more flexible and potentially more discriminating than other efficiency criteria based on stochastic dominance.

Results of an empirical test of the interval approach to the measurement of decision maker preferences demonstrate its value. They show that it leads to a lower probability of eliminating a decision

maker's preferred choice from the efficient set of choices than that realized with a single-valued utility function and that permits the identification of efficient sets smaller than those associated with the criteria of first and second degree stochastic dominance. The results also show how the precision with which preferences are measured can be varied under this approach.

The identification of preferred choices is the primary objective of any decision analysis. The final methodological contribution of this study is the formulation of a generalized risk efficient Monte Carlo programming model (GREMP), which combines random search procedures, Monte Carlo simulation, and evaluation by the criterion of stochastic dominance with respect to a function in a single analytical framework for the identification of preferred choices. This model is both flexible and computationally efficient, and it is well suited for the analysis of a wide range of practical decision problems. The use of Monte Carlo programming procedures to generate alternative strategies for consideration facilitates the introduction of flexibility into the definition of a management strategy, since strategies can be defined by feedback control rules as well as by specific values of control variables over the entire planning horizon. The incorporation of Monte Carlo simulation techniques into the model facilitates the more realistic representation of the complex processes by which the control variable levels determined by a particular management strategy and a set of stochastic environmental factors interact to determine the properties of the distribution of outcomes associated with any choice. Finally, evaluation by the criterion of stochastic dominance with respect to a

function permits the ordering of choices in a manner fully consistent with the expected utility hypothesis without requiring that restrictive assumptions be made about decision maker preferences or the form of system output distributions.

The methodological tools developed in this study are of considerable value. It should be noted, however, that they are not intended for use in all decision situations. They allow considerable flexibility concerning the degree of detail incorporated into any decision analysis, but they are intended primarily for application on a computer and can be expensive to use. They do not replace budgeting techniques or mathematical programming but supplement these and other existing methodological tools. As is true of any aid in the decision process, the procedures developed here should be employed only when the benefits associated with their use exceed the possible added costs. It should also be noted that these procedures are not all that is required to successfully resolve a practical decision problem. Rather, they represent a set of tools which facilitates some aspects of the decision process--a process which involves a wide range of activities.

7.2 Empirical Findings

This study focuses primarily on methodological rather than empirical issues. Several empirical findings associated with the illustrative applications of the procedures developed in this study are worthy of note, however.

As a test of the interval approach to the measurement of decision maker preferences, preference measurements were made for seventeen farmer participants in a marketing extension workshop. As reported in Section 4.8

of Chapter IV, the resulting preference measurements provide rather strong evidence that many decision makers exhibit both risk loving and risk averse behavior. The interval measurements for thirteen of the seventeen respondents included negative as well as positive values. The results also showed that strictly decreasing absolute risk aversion functions are not as common as is often suggested. Though certainly not generalizable to a larger population, these findings do call to question the wide acceptance of the proposition that absolute risk aversion functions tend to be positive valued and strictly decreasing, and so suggest that further research in this area may be warranted.

The procedures developed in this study are applied in the analysis of two related examples concerned with choices made in the operation of a cash grain farm. The first example focuses on land rental and production planning decisions when prices, yields, and time available for fieldwork are uncertain. In the second example, these same decisions are considered in conjunction with the choice of a flexible marketing strategy involving cash sales at harvest and forward contracting. The results of both applications demonstrate the flexibility of the procedures developed in this study. They also indicate that decision maker preferences have an important impact on both the selection of a cropping plan and on the choice of an appropriate marketing strategy. Finally, they show that the degree of interdependence between production and marketing strategies is not particularly great, which suggests that it may be possible to analyze these two components of a management strategy separately in some situations.

7.3 Implications for Future Research

The techniques developed in this study can be of use in the analysis of a diverse range of practical decision problems. Clearly more work can be done, for example, on the analysis of alternative production and marketing strategies for agricultural firms. Of particular value would be attempts to develop flexible marketing strategies which are suitable for a broad range of decision makers and decision situations. Another potentially important application of these techniques is in the analysis of alternative pest management strategies. Such strategies are, in effect, feedback control rules designed to direct actions in complex decision situations characterized by a considerable degree of uncertainty, and the GREMP model, with its combination of random search and simulation, is well suited for the identification and evaluation rules of this type.

The methodological tools developed in this study can also be of use in the analysis of major investment-disinvestment decisions, both private and public. They permit the explicit consideration of flexibility in the analysis of such choices--a factor which, as Massé (1962) notes, is of critical importance when major resource commitments are to be made in an uncertain environment. At a still higher level of aggregation, these tools can also be employed in the analysis of policy decisions having outcomes which are strongly influenced by stochastic factors in the environment in which they are implemented.

The techniques developed in this study greatly facilitate the application of decision theory based on the expected utility hypothesis in the analysis of practical decision problems. They have considerable potential value, but they can be made more effective if they are refined

still further. There is a need, for example, for further testing of the procedures used to implement the interval approach to the measurement of decision maker preferences. Experiments should be conducted to identify measurement scales which allow preferences to be adequately represented in a wide range of decision situations, and alternative modes of questioning should be evaluated. Will a measurement grid which works well in the analysis of choices having outcomes which are concentrated around a certain value be adequate when preferences are to be measured over a much broader range of outcomes? Over how wide a range of systems output levels can absolute risk aversion be assumed to be constant? Does the range depend on the level of system output? In the neighborhood of how many system output levels should preferences be measured? Upon how many choices should each measurement be based? These are but five of the many questions which need to be answered.

More research is also needed on the representation of preference relationships which depend on more than one system output variable and on the development of multivariate stochastic dominance criteria. Though some work has been done in the latter area by Levy (1973), Levy and Paroush (1974a, b) and Kihlstrom and Mirmon (1974), further research is needed. Particularly valuable would be an extension of stochastic dominance with respect to a function to the multivariate case.

Additional refinements are also needed in the GREMP model. It may be possible, for example, to increase the efficiency of this procedure by incorporating learning rules which lend at least partial direction to the random search. Such rules could be applied at the end of each 100 iterations of the model and might have the effect of reducing the

range of values for control variables over which the search is to be made in subsequent iterations. A second type of rule might cause the search to be stopped if no new strategies have entered the efficient set for a certain number of iterations.

Finally, it must be noted that methodological advances alone will not eliminate all the difficulties encountered in the applied analysis of decisions made under uncertainty. The two applications of the techniques developed in this study point to the importance of and the need for an improved information base in most decision situations. This need is particularly strong with regard to decisions made by agricultural producers who face so many different types of uncertainty. Efforts must be made to supply producers with probabilistic price forecasts and to teach them to use such information effectively. Frequently, all the information needed to make forecasts in probabilistic terms is readily available to the agencies or individuals who predict future price levels, but in most cases price forecasts simply state a most likely value, or at best, an interval within which the price is expected to fall. Similarly, more complete information is also needed about how yields are affected by timeliness and by stochastic factors in the environment. More consideration needs to be given to such factors in the design of agronomic experiments.

The information base for the analysis of decisions made under uncertainty could also be improved by research designed to identify systematic relationships between levels of absolute risk aversion and selected decision maker characteristics. Such information could be of use to policy analysts who wish to consider the impacts of uncertainty

on the choices made by representative firms. It could also be of use in situations when the importance of a choice to be made does not warrant the expenditure of sources required to construct an interval measurement of the decision maker's preferences.

In conclusion, then, the methodological tools developed in this study can, in their present form, be employed in the analysis of a wide range of practical decision problems. They represent an important improvement in the set of techniques available for the applied analysis of decisions made under uncertainty. Further efforts are needed, however, to make these tools both easier to use and more effective.

APPENDICES

APPENDIX A

A GENERALIZED MULTIVARIATE PROCESS GENERATOR

A.1 Introduction

The analysis of decisions made under uncertainty requires that information on the probability distributions of system output variables be determined for each alternative strategy being considered. In most cases the properties of such distributions depend on the controllable system input levels which define any particular strategy and on the probability distributions of stochastic environmental factors which cannot be controlled by the decision maker. When there is only one stochastic environmental factor or when the relationship between controllable and exogenous system inputs and system outputs is relatively simple, the properties of system output distributions can be derived analytically.¹ When several stochastic factors having probability distributions from different families must be considered or when input-output relationships are not of a convenient form, however, analytical techniques cannot be used to determine the properties of system output distributions. In such instances, Monte Carlo sampling procedures and numerical simulation techniques are frequently used to obtain the necessary information. Sample states of nature are defined by selecting

¹See Anderson and Doran (1978) for an excellent review of the conditions under which information on such probability distributions can be determined analytically.

values for each exogenous system input in a pseudorandom fashion. Numerical simulation techniques are then used to determine the system output levels associated with a particular strategy for each state of nature. The resulting values constitute a random sample from the probability distribution of system outputs. They can be used to calculate sample moments of the underlying distribution, or they can be used to construct an approximate representation of the cumulative distribution function of the system output variable.

A process generator is a procedure, usually programmed for implementation on a computer, which generates pseudorandom sample observations from a specified probability distribution.¹ As such, process generators are a basic building block in the procedure described above. Process generators have been developed for a wide range of standard univariate probability distributions including the discrete and continuous uniform, exponential, gamma, beta, chi-square, normal, lognormal, geometric, binomial, hypergeometric, and Poisson (Naylor, et al., 1966; Schmidt and Taylor, 1970; Newmann and Odell, 1971). In determining the properties of a distribution of systems outputs, the use of univariate process generators is appropriate if all underlying stochastic factors can be assumed to be independently distributed. When this assumption cannot be made, a multivariate process generator is required.

Process generators have been developed for several multivariate distributions, most notably the multivariate normal and the Wishart

¹The term "pseudorandom" refers to the fact that, although values generated by a process generator have all the properties of a random sample, they are actually generated in a deterministic fashion.

distributions (Naylor, et al., 1966; Newmann and Odell, 1971). In many cases, however, the properties of the joint distribution of a particular set of stochastic environmental factors may not be adequately approximated by either of these distributions. There is a need, then, for multivariate process generators which are more flexible than those currently available.

A major contribution of this study is the development of a workable procedure for the generation of random variates from multivariate probability distributions with non-normal marginal distributions. The formulation and implementation of this procedure is the primary focus of this appendix. Before introducing the generalized multivariate process generator, however, basic concepts related to the generation of random variates are first reviewed, and several commonly used univariate process generators are presented. The generation of random variates from the multivariate normal distribution is then discussed, since this procedure is used in the more general process generator developed in this study. Finally, the algorithm for the generation of random variates from multivariate distributions with specified marginal distributions and correlation matrix is presented along with an explanation and listing of the computer program used to implement it.¹ An empirical example is also presented to demonstrate the efficacy of this procedure.

¹Though developed independently, this procedure is similar in many respects to methods described in an unpublished paper by Coleman and Saipé (1976), which describes several procedures for the generation of sample observations from bivariate distributions.

A.2 Basic Procedures for the Generation of Random Variates

Several approaches have been developed for the generation of random variates. They include the inverse transformation method, the rejection method, and the composition method, all of which are described in Naylor et al. (1966). Because it is the most commonly used and because it serves as the basis for the algorithm to be discussed below, only the inverse transformation method will be reviewed here.

The cumulative distribution function, $F(x)$, of the continuous random variable, x , is defined over the interval $(0,1)$. Associated with each value of x , then, is a value r lying on the interval $(0,1)$ such that

$$r = F(x) \quad \text{A.1}$$

Similarly, if the inverse of $F(x)$ can be determined, the following relationship will hold:

$$x = F^{-1}(r) \quad \text{A.2}$$

In this case any particular value of r uniquely determines a value of x if $F^{-1}(r)$ is a continuous, monotonically increasing function. By generating a set of uniformly distributed random variables lying on the interval $(0,1)$ and calculating the corresponding values of x determined by equation A.2, a set of sample observations from the probability distribution of x is generated. This is the inverse transformation method for generating random variates.

Following Manetsch and Park (1977b) the validity of the inverse transformation method can be demonstrated in the following manner. Let

$$r = F(x) \quad \text{A.3}$$

where r is a uniformly distributed random variable defined on the interval $(0,1)$. By the definition of a cumulative distribution function

$$F(x_0) = P\{x \leq x_0\} = r_0 \quad A.4$$

where x_0 is a specific value of x and r_0 is the corresponding value of r . Since r is uniformly distributed

$$G(r_0) = P\{r \leq r_0\} = r_0^1 \quad A.5$$

Since $F(x_0)$ and r_0 are equal, Equation A.5 implies that

$$P\{r \leq F(x_0)\} = F(x_0) \quad A.6$$

and since $F(x)$ is assumed to be a monotonically increasing function in x , it also follows that

$$P\{F^{-1}(r) \leq x_0\} = F(x_0) = P\{x \leq x_0\} \quad A.7$$

Therefore, the distribution of the random variable generated by the inverse transformation method is identified to that of the random variable x .

Application of the inverse transformation method is shown graphically in Figure A.1. A cumulative distribution function $F(x)$ is drawn on the left, and its inverse is drawn on the right. In the particular example shown the randomly selected value of R is .5, and the associated value of x is 6.3.

The inverse transformation method is the basis for commonly used procedures for generating random variates from several standard probability

¹For uniformly distributed random variables

$$g(r_0) = \begin{cases} r_0 & 0 \leq r_0 \leq 1 \\ 0 & \text{elsewhere} \end{cases}$$

and

$$G(r_0) = \int_0^{r_0} g(y) dy = \int_0^{r_0} y dy = r_0$$

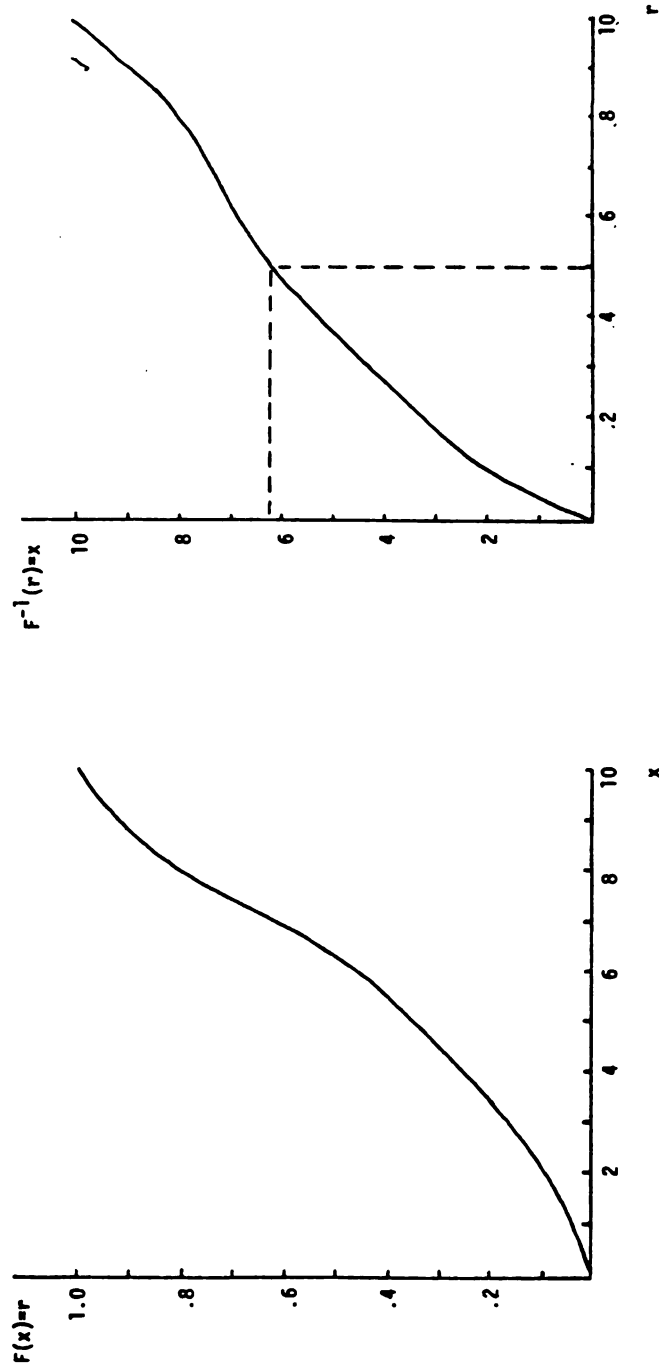


Figure A.1 Generation of Random Variates by the Inverse Transformation Method

distributions. Most notably, it is the basis for process generators for exponential, gamma and beta random variables. Procedures for the generation of variates from each of these distributions are presented below with the computer programs which implement them.¹

A.2.1 An Exponential Process Generator

The density function of an exponential random variable is given by the expression:

$$f(x) = \begin{cases} \alpha e^{-\alpha x} & 0 \leq x \\ 0 & \text{otherwise.}^2 \end{cases} \quad \text{A.8}$$

This function can be integrated to obtain the cumulative distribution function, $F(x)$

$$F(x) = \int_0^x \alpha e^{-\alpha x} dx = 1 - e^{-\alpha x} = r. \quad \text{A.9}$$

The inverse of $F(x)$ is given by the expression

$$F^{-1}(r) = x = -\frac{1}{\alpha} \ln(1-r). \quad \text{A.10}$$

If r is a uniformly distributed random variable on the interval $(0,1)$ this expression reduces to

$$x = -\frac{1}{\alpha} \ln(r). \quad \text{A.11}$$

By generating values of r at random and calculating the associated values of x using Equation A.11, sample observations from an exponential distribution with parameter α can be generated.

A computer program which implements this procedure is listed in Figure A.2. Three parameter values must be supplied by the user: ND, BL, and XMEAN. ND is the number of variates to be generated, BL is

¹These programs can easily be made into subroutines and incorporated into larger programs.

²The mean and variance of this distribution are $1/\alpha$ and $1/\alpha^2$ respectively.


```

PROGRAM EXPRV(INPUT,OUTPUT,TAPE5=INPUT,TAPE6=OUTPUT)
READ(5,100) ND,BL,XMEAN
A=XMEAN-BL
DO 5 I=1,ND
R=RANF(-1)
X=-A*ALOG(R)+BL
WRITE(6,200) X
100 FORMAT(I5,2F10.2)
200 FORMAT(F10.2)
END

```

Figure A.2 A Program for Generating Exponential Random Variates

the lowest value the random variable can take,¹ and XMEAN is the expected value of the distribution.

A.2.2 A Gamma Process Generator

The density function of a gamma distribution is

$$f(x) = \begin{cases} \frac{\alpha^k x^{(k-1)} e^{-\alpha x}}{(k-1)!} & 0 \leq x \\ 0 & \text{otherwise}^2 \end{cases} \quad \text{A.12}$$

An analytic expression for the cumulative distribution function of this distribution does not exist. It can be shown, however, that the sum of k exponentially distributed random variables, each having an expected value of $1/\alpha$, is a gamma random variable with a density function identical to that given in Equation A.12 (Schmidt and Taylor, 1970, p. 265).

Random variates from gamma distributions for which the parameter k is an integer, then, can be generated by summing k variates drawn from an appropriate exponential distribution.

A computer program which generates gamma random variates by this method is given in Figure A.3. In this case four parameter values must be supplied by the user: ND, BL, XMEAN, and K. ND, BL, and XMEAN are, again, the number of sample observations to be generated, the lower bound of the distribution, and the expected value of the distribution.

¹In the deviation above, BL was assumed to be zero, but this need not be the case.

²The mean and variance of this distribution are k/α and k/α^2 respectively.

```

PROGRAM GAMMA(INPUT,OUTPUT,TAPE5=INPUT,TAPE6=OUTPUT)
READ(5,100) ND,RL,XMEAN,K
A=(XMEAN/RL-RL)/FLOAT(K)
DO 10 I=1,ND
  V=RL
  DO 5 J=1,K
    R=RANF(-1)
    E=-A+ALOG(R)
    Y=Y+E
  5 WRITE(6,200) Y
  100 FORMAT(15,2F10.2,15)
  200 FORMAT(F10.2)
  END

```

Figure A.3 . A Program for Generating Gamma
Random Variates

K is the second parameter of the gamma distribution.¹ The parameter K cannot always be treated as an integer. Though not reviewed here, a procedure for generating variates from gamma distribution with non-integer parameters is described elsewhere by Phillips and Beightler (1972).

A.2.3 A Beta Process Generator

The density function of the beta distribution is

$$f(x) = \frac{\Gamma(\alpha+\beta)x^{\alpha-1}(1-x)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} \quad 0 \leq x \leq 1 \quad \text{A.13}$$

$$0 \quad \text{elsewhere}^2$$

Again, an analytic expression for the cumulative distribution function cannot be derived. As Naylor et al. (1966) note, however, the random variable defined by the expression

$$x = x_1/(x_1+x_2), \quad \text{A.14}$$

where x_1 and x_2 are both gamma random variables with identical values of α and values of k such that $k_1 = \alpha$ and $k_2 = \beta$, has a beta distribution with parameters α and β . If α and β are integers, then, beta random variates can be generated using an extension of the procedures

¹The parameter k can be determined by solving the following expression

$$k = \frac{(\mu - BL)^2}{\alpha^2}$$

where μ and α^2 are the mean and variance of the distribution being modelled and BL is the lower bound of that distribution. The program computes the value of the parameter α automatically.

²The mean and variance of this distribution are

$$\mu = \alpha/(\alpha+\beta)$$

and

$$\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$$

developed above. Two gamma variates from appropriate distributions are generated and a beta variate is defined according to Equation A.14.¹

A computer program for generating beta variates is given in Figure A.4. The user must supply values for five parameters: ND, BL, BU, K1, and K2. ND is the number of variates to be generated. BL and BU are the upper and lower bound values which the random variable can take. Though set at zero and one in the derivation above, they can be set at any values. K1 and K2 are the two shape parameters of the beta distribution. Their values are determined by solving the following two equations

$$K2 = \frac{\gamma^3 - 2\gamma^2 + (1+\delta)\gamma - \delta^2}{\delta^2} \quad A.15$$

$$\text{and } K1 = \frac{\gamma K2}{1-\gamma} \quad A.16$$

where γ and δ^2 are the mean and variance of the distribution once it has been normalized so that all values be on the interval (0,1).²

A.2.4 A Generalized Univariate Process Generator

Each of the process generators above is based on a simple application of the inverse transformation method to the exponential distribution. For many distributions, however, the inverse of the cumulative

¹When a beta distribution has non-integer parameters, the gamma variates required in Equation A.14 can be generated using the procedure developed by Phillips and Beightler (1972).

²More formally,

$$\gamma = (\mu - BL) / (BU - BL)$$

and

$$\delta^2 = \gamma^2 / (BU - BL)^2$$

where μ and γ^2 are the mean and variance of the actual distribution to be modelled.

```

PROGRAM BETA(INPUT,OUTPUT,TAPE5=INPUT,TAPE6=OUTPUT)
READ(5,100) ND,PL,RU,K1,K2
DO 15 I=1,ND
  G1=0.0
  G2=0.0
  DO 5 J=1,K1
    R=RANF(-1)
    F=-ALOG(F)
    G1=G1+F
  5
  DO 10 J=1,K2
    R=RANF(-1)
    F=-ALOG(F)
    G2=G2+F
  10
  X=(G1/(G1+G2))*(RU-PL)+PL
  15
  WRITE(6,200) X
  100 FORMAT(IE,2F10.2,2IF)
  200 FORMAT(F10.2)
END

```

Figure A.4 A Program for Generating Beta Random Variates

distribution function cannot be derived analytically, and simple analytical links to other distributions for which such inverses can be calculated may not exist. Most notably, this is true for the normal distribution. In the case of many empirical probability distributions, the exact functional form of the cumulative distribution function may not even be known. Nevertheless, the inverse transformation method can still be used in such instances. Values of the cumulative distribution function $F(x)$ at specified levels of the random variable x can be used to construct a rough approximation of the entire function by linearly interpolating between known points.¹ In Figure A.5, for example, six points on a cumulative distribution function are specified: the upper and lower bounds, and three intermediate points which divide the cumulative into quartiles. Given such a diagrammatic representation, the value of the random variable x associated with any randomly selected probability level is easily determined. As shown in Figure A.5, for example, the value of x corresponding to a probability level of .2 is 44.

A table look-up function (Llewellyn, 1965) is the equivalent of such a diagram on a digital computer. The table look-up function TABEX, a listing of which is given in Figure A.6, has access to an array of X values, ARG, and an array of corresponding probability levels, VAL, each array having K elements. Given an argument value,

¹Points in the cumulative distribution function can be determined analytically, by numerical integration of a known probability density function, or they can simply be sample observations from an empirical distribution. As Schlaeffer (1959) notes sample observation arranged in order are reasonable estimates of the fractiles of the underlying distribution.

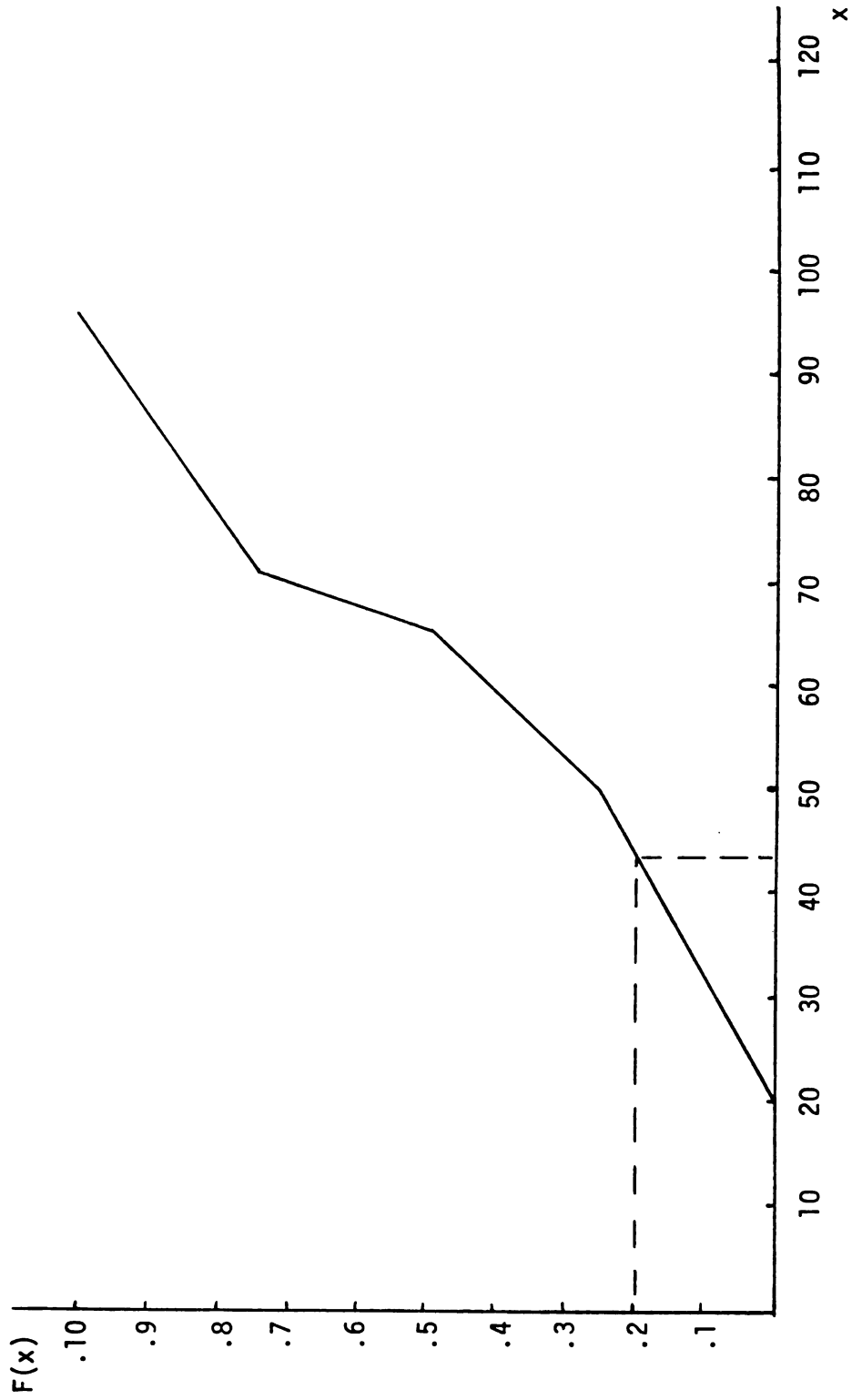


Figure A.5 An Approximate Representation of a Cumulative Distribution Function Based on Six Known Points


```

      FUNCTION TAPEX(VAL,ARG,DUMMY,K)
      DIMENSION VAL(1),ARG(1)
      GO TO 1 J=2,N
      IF(DUMMY.GT.ARG(J)) GO TO 1
2  TAPEX=(DUMMY-ARG(J-1))*(VAL(J)-VAL(J-1))/
1  (ARG(J)-ARG(J-1))+VAL(J-1)
      RETURN
1  CONTINUE
      J=K
      GO TO 2
      END

```

Source: Llewellyn (1965, p. 4-22).

Figure A.6 A Table Look-up Function Subprogram

DUMMY, this function subprogram computes a value of the function defined by ARG and VAL. If x is a random variable and if R is a variable lying on the interval $(0,1)$ which corresponds to a value of $F(x)$, then the FORTRAN statement

$R = \text{TABEX}(\text{VAL}, \text{ARG}, X, K)$ A.17

calculates the probability level on the cumulative distribution function corresponding to any specified value of X . One of the interesting properties of this particular table look-up function is that by simply switching the positions of VAL and ARG in the calling statement, values of the inverse function can be calculated. The FORTRAN statement

$X = \text{TABEX}(\text{ARG}, \text{VAL}, R, K)$ A. 18

for example, calculates the value of X which corresponds to any specified probability level, R .

A computer program which uses the inverse transformation method to generate variates from a probability distribution defined by a specified set of points on the cumulative distribution function is given in Figure A.7. The user must supply values of ND, the number of variates to be generated, and K, the number of points on the cumulative distribution function to be specified explicitly. Paired values of the random variable and the cumulative distribution function are then read into arrays ARG and VAL. These values must be arranged by the user in ascending order.

The procedures developed in this section are of interest in themselves, since they are all useful tools for Monte Carlo sampling. More important in relation to the primary purpose of this appendix, however, they are used extensively in the generalized procedure for the generation of variates from multivariate distribution.

```

      PROGRAM RVGEN(INPUT,OUTPUT,TAPE5=INPUT,TAPE6=OUTPUT)
      DIMENSION ARG(50),VAL(50)
      C READ NUMBER OF VARIATES TO BE GENERATED
      READ(5,100) ND
      C READ NUMBER OF POINTS ON CUMULATIVE DISTRIBUTION FUNCTION
      C TO BE SPECIFIED. K CANNOT EXCEED 50.
      READ(5,100) K
      C READ VALUES OF VAL AND ARG IN ASCENDING ORDER.
      DO 5 I=1,K
      5  READ(5,101) VAL(I),ARG(I)
      C GENERATE RANDOM VARIATES
      DO 10 I=1,ND
      R=RNPF(-1)
      X=TAPEX(ARG,VAL,P,K)
      10  WRITE(6,200) X
      100  FORMAT(I5)
      101  FORMAT(2F10.2)
      200  FORMAT(F10.2)
      END

      FUNCTION TAPEX(VAL,ARG,DUMMY,K)
      DIMENSION VAL(1),ARG(1)
      DO 1 J=2,K
      IF(DUMMY.GT.ARG(J)) GO TO 1
      2  TAPEX=(DUMMY-ARG(J-1))*(VAL(J)-VAL(J-1))/
      1 (ARG(J)-ARG(J-1))+VAL(J-1)
      RETURN
      1  CONTINUE
      J=K
      GO TO 2
      END

```

Figure A.7 A Generalized Univariate Process Generator

A.3 The Generation of Variates from the Multivariate Normal Distribution

As stated above, the multivariate normal distribution is one of the few multivariate probability distributions from which a workable process generator has been developed. Each element of the vector of random variables, x , of which a multivariate distribution is comprised is normally distributed with specified mean and variance. The multivariate distribution is fully described by the vector of expected values for each of its marginals, μ , and by a positive definite, symmetrical variance-covariance matrix, Ω , which is defined by the following expression.

$$\Omega = E[(x-\mu) \cdot (x-\mu)] \quad A.19$$

If the elements of x are not correlated, the off-diagonal elements of this matrix will equal zero and each variate in x can be generated independently using procedures such as those outlined in the preceding section. When correlations are present, however, this approach is not satisfactory.

Naylor, et al. (1966) describe a procedure for the generation of variates from the multivariate normal distribution which is based on a theorem proved by T. W. Anderson (1959). That theorem states that if z is a vector of independent standard normal variates, there exists a unique lower triangular matrix, C , such that

$$x = Cz + \mu \quad A.20$$

It follows directly that the variance-covariance matrix of $(x-\mu)$ --which is also the variance-covariance matrix of x , since μ is a vector of constants--is defined by the expression CC' . Therefore,

$$\Omega = CC' \quad A.21$$

The "square root method" can be used to derive a set of recursive formulas for computing the elements of C from those of Ω .

Once the elements of C have been calculated, a vector of independent standard normal variates, z , can be generated, and the vector x can be found using Equation A.20. To generate a large number of sample vectors from a given multivariate distribution, the elements of C need to be calculated only once. The final two steps described above are then repeated for the generation of each vector of variates.

A computer program which implements this procedure is listed in Figure A.8. The user of this program must specify ND and MN , the number of sample vectors to be generated and the number of elements in each vector. He must also specify the mean and variance of each of the MN marginal distributions. Finally, the off-diagonal elements of the correlation matrix which are non-zero must be specified.¹ The variable IND is set to a non-zero value when the last correlation coefficient is read. This program can be used to model multivariate normal distributions with up to fifty elements. More information about its structure is provided on comment cards included in the listing.

A.4 A Generalized Multivariate Process Generator

In this section a generalized procedure for the generation of sample vectors from multivariate probability distributions is described.

¹For any pair of random variables x and y , then correlation coefficient ρ_{xy} is defined by the expression

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

where σ_{xy} is the covariance between x and y and σ_x and σ_y are the standard deviation of the two random variables.

```

PROGRAM MVNOR(INPUT,OUTPUT,TAPE5=INPUT,TAPE6=OUTPUT)
COMMON /BLUCK1/ C(50,50),COR(50,50),PAR(50,50),V(50),MN
DIMENSION X(50),XMEAN(50),Z(50)
C READ NUMBER OF SAMPLE VECTORS TO BE GENERATED AND NUMBER
C OF ELEMENTS IN EACH VECTOR.
READ(5,100) NI,MN
C READ MEAN AND VARIANCE OF EACH MARGINAL DISTRIBUTION
DO 5 I=1,MN
5 READ(5,101) XMEAN(I),V(I)
C CONSTRUCT CORRELATION MATRIX
C FIRST SET ALL ELEMENTS EQUAL TO ZERO.
DO 10 I=1,MN
DO 10 J=1,MN
COR(I,J)=0.0
IF(I.EQ.J) COR(I,J)=1.0
10 CONTINUE
C NEXT READ ALL NON-ZERO ELEMENTS IN LOWER TRIANGULAR PORTION
C OF MATRIX SET IND EQUAL TO A NON-ZERO VALUE WHEN THE LAST
C ELEMENT IS TO BE READ
15 READ(5,102) I,J,COR(I,J),IND
C THIS STATEMENT SETS NON-ZERO VALUES IN THE UPPER TRIANGULAR
C PORTION OF THE MATRIX
COR(J,I)=COR(I,J)
IF(IND.EQ.0) GO TO 15
C CONSTRUCT LOWER TRIANGULAR MATRIX C.
CALC COEF
C BEGIN TO GENERATE SAMPLE VECTORS.
DO 35 I=1,NI
C GENERATE MN INDEPENDENTLY DISTRIBUTED STANDARD NORMAL VARIATES.
DO 20 J=1,MN
R1=RNRF(-1)
R2=RNRF(-1)
C THIS TECHNIQUE FOR THE GENERATION OF NORMAL VARIATES IS FROM
C MANETSCH AND PARK(1977B)
20 Z(J)=(-2.0*ALOG(R1))**.5*COS(6.2831*R2)
C GENERATE VARIATES FROM DESIRED MULTIVARIATE NORMAL DISTRIBUTION.
DO 30 J=1,MN
X(J)=XMEAN(J)
DO 25 K=1,MN
25 X(J)=X(J)+C(J,K)*Z(K)
30 CONTINUE
35 WRITE(6,200) (X(J),J=1,MN)
100 FORMAT(2I5)
101 FORMAT(2F10.2)
102 FORMAT(2I5,F10.2,I5)
200 FORMAT(10F8.2)
END

SUBROUTINE COEF
C THIS SUBROUTINE CONSTRUCTS THE LOWER TRIANGULAR MATRIX C.
C IT WAS WRITTEN BY R. NEAL PETERSON
COMMON /BLUCK1/ C(50,50),COR(50,50),PAR(50,50),V(50),MN
DO 10 I=1,MN
DO 10 J=1,MN
C(I,J)=0.0
10 DO 20 I=1,MN
20 PAR(I,1)=COR(I,1)
IF(MN.LT.2) GO TO 70
DO 60 J=2,MN
JM1=J-1
SUM=0.
DO 30 K=1,JM1
30 SUM=SUM+PAR(J,K)**2
PAR(J,J)=SQRT(1-SUM)
IF(J.EQ.MN) GO TO 70
JP1=J+1
DO 50 I=JP1,MN
50 PAR(I,J)=0.
IF(PAR(J,J).EQ.0.) GO TO 50
SUM=0.
DO 40 K=1,JM1
40 SUM=SUM+PAR(I,K)*PAR(J,K)
PAR(I,J)=(COR(I,J)-SUM)/PAR(J,J)
50 CONTINUE
60 CONTINUE
70 DO 80 I=1,MN
DO 80 J=1,I
80 C(I,J)=PAR(I,J)*SQRT(V(I))
RETURN
END

```

Figure A.8 A Program for Generating Multivariate Normal Random Variates

This procedure can be used to generate random variates from multivariate distributions having up to fifty marginal distributions. Marginal distributions can be of any form, and they need not all belong to the same family of distributions. All that is required is that enough information be available to construct the cumulative distribution function of each marginal and to specify the correlation coefficient between each pair of random variables in the distribution.¹ The only condition on the matrix of correlation coefficients is that it be positive definite and symmetrical, as required for internal consistency.

The generalized multivariate process generator described below is based on the premise that correlations between the elements of a marginal distribution are, to a large extent, preserved as the elements undergo successive inverse transformations. The degree to which correlations are actually preserved under such transformations is difficult to determine analytically in all but the simplest of cases. Levels of correlation are not expected to be maintained exactly. It is expected, however, that high levels of correlation between elements of the initial distribution will lead to high levels of correlation between corresponding elements of the transformed distribution. All experience to date, some of which is reported below, strongly supports this basic premise.

The procedure for generating sample observations from a generalized multivariate probability distribution is relatively simple. Once the

¹As Coleman and Saipe (1976) observe, the higher order product moments of a multivariate probability distribution may have an important impact on its character. These moments are difficult to estimate reliably and are rarely considered explicitly in the description of multivariate processes. They cannot be considered in the process developed here.

cumulative distribution function of each marginal and a correlation matrix have been specified for the multivariate distribution to be modelled, a simple vector, z is generated from a multivariate normal distribution having a correlation matrix identical to that specified. For the sake of convenience all the marginals of the multivariate normal distribution are standard normal. Each element of the vector z is then transformed to a uniformly distributed random variable defined on the interval $(0,1)$. This is done by simply finding the value on the cumulative distribution of a standard normal random variable associated with each element of z .¹ The resultant vector, u , is a sample vector from a multivariate distribution, all marginals of which are uniform distributions defined in the interval $(0,1)$. The correlation matrix of that distribution is not known exactly, but it is expected that it can be reasonably well approximated by the correlation matrix originally specified. Finally, each element of the vector u is transformed by the inverse transformation method to a sample observation from the corresponding marginal distribution of the multivariate distribution being modelled. The resultant vector, x , can be viewed as a sample observation from a multivariate distribution having marginals identical to that being modelled and a correlation matrix which, though not necessarily identical to that originally specified, is expected to be a reasonably good approximation of that matrix. For most purposes this procedure should provide an adequate representation of the distribution being modelled. The sample properties for an example are discussed below.

¹The fact that the transformed variable is uniformly distributed follows directly from the proof of the validity of the inverse transformation method given in Section A.2.

A computer program which implements this procedure is listed in Figure A.9. Several types of inputs must be supplied by the user of MVGEN. ND and MN are the number of sample vectors to be generated and the number of elements in each vector. MN cannot exceed fifty. K is the number of data points to be specified for the construction of the cumulative distribution function of each marginal distribution. It can take a value as high as 100. K paired values of the random variable and cumulative distribution function of each marginal must be supplied by the user. These values, which must be arranged in ascending order, are read into the two dimensional arrays ARG and VAL. Finally, non-zero off diagonal elements of the correlation matrix must be supplied by the user. These can be read in any order. The variable IND is set to a non-zero value when the last correlation coefficient is read.

The structure of the program closely follows the outline of the procedure described above. After all necessary variables are initialized, sample vectors are generated sequentially by first generating a sample vector from a multivariate normal distribution and then transforming the elements of that vector first to variates of a uniform distribution and finally to variates from each marginal of the multivariate distribution being modelled. More information about particular aspects of the program is given in comment cards.

Program MVGEN is designed to permit the user to specify marginal distribution of any form. When all the marginals of a particular multivariate distribution can be adequately described by one or more standard Probability distributions, it may be more convenient and more efficient Computationally to modify MVGEN so that only the parameters of each

```

      PROGRAM MVGEN(INPUT,OUTPUT,TAP5=INPUT,TAP6=OUTPUT)
      COMMON /BLOCK1/ C(50,50),COR(50,50),PAR(50,50),MN
      COMMON /BLOCK2/ APC(50,100),VAL(50,100)
      DIMENSION RV(50),U(50),E(50),VALN(41),ARGN(41)
      DIMENSION Y(50)
C THE NEXT TWO DATA STATEMENTS ASSIGN VALUES TO VARIABLES
C USED TO CONSTRUCT A TABLE LOOK UP REPRESENTATION OF THE
C INVERSE OF THE CUMULATIVE DISTRIBUTION FUNCTION OF A
C STANDARD NORMAL RANDOM VARIABLE.
      DATA KN/41/
      DATA ARGN/ -3.5, -1.96, -1.645, -1.439, -1.281, -1.150, -1.037, -.925,
1-.841, -.758, -.674, -.598, -.524, -.454, -.386, -.312, -.253, -.189,
2-.128, -.068, 0.0, .054, .126, .187, .253, .312, .386, .454, .524, .598,
3.674, .755, .841, .925, 1.037, 1.150, 1.281, 1.439, 1.645, 1.960, 3.5/
C READ THE NUMBER OF SAMPLE VECTORS TO BE GENERATED, ND, AND
C THE NUMBER OF ELEMENTS IN EACH VECTOR, MN.
      READ(5,100) ND,MN
C READ THE NUMBER OF PRINTS TO BE DEFINED ON THE CUMULATIVE
C DISTRIBUTION OF EACH MARGINAL, K.
      READ(5,101) K
C READ DATA POINTS FOR EACH MARGINAL.
      DO 5 I=1,MN
      DO 5 J=1,K
5 READ(5,102) ARG(I,J),VAL(I,J)
C SET ALL OFF-DIAGONAL ELEMENTS OF THE CORRELATION MATRIX TO ZERO.
      DO 10 I=1,MN
      DO 10 J=1,MN
      COR(I,J)=0.0
      IF(I.EQ.J) COR(I,J)=1.0
10 CONTINUE
C READ ALL NON-ZERO CORRELATION COEFFICIENTS, SET IND EQUAL
C TO A NON-ZERO VALUE WHEN THE LAST CORRELATION COEFFICIENT IS
C TO BE READ.
15 READ(5,103) I,J,COR(I,J),IND
      COR(J,I)=COR(I,J)
      IF(IND.EQ.0) GO TO 15
C CONSTRUCT LOWER TRIANGULAR MATRIX C.
      CALL COEF
      VALN(1)=0.0
      DO 20 I=2,KN
      J=I-1
20 VALN(I)=VALN(J)+.025
C BEGIN TO GENERATE RANDOM VARIATES.
      DO 50 KK=1,ND
C CALL NORVEC, WHICH RETURNS A VECTOR OF INDEPENDENTLY DISTRIBUTED
C STANDARD NORMAL RANDOM VARIATES.
      CALL NORVEC(MN,E)
C CONSTRUCT SAMPLE VECTOR FROM A MULTIVARIATE NORMAL DISTRIBUTION.
      DO 30 I=1,MN
      RV(I)=0.0
      DO 25 J=1,MN
25 RV(I)=RV(I)+C(I,J)*E(J)
30 CONTINUE
C TRANSFORM THE VECTOR OF MULTIVARIATE NORMAL RANDOM VARIABLES
C INTO A VECTOR OF UNIFORMLY DISTRIBUTED RANDOM VARIABLES.
      DO 35 I=1,MN
35 U(I)=TAREX(VALN,ARGN,RV(I),KN)
C TRANSFORM THE UNIFORMLY DISTRIBUTED RANDOM VARIABLES INTO
C RANDOM VARIABLES FROM THE SPECIFIED BETA DISTRIBUTIONS.
      DO 40 I=1,MN
40 Y(I)=TASEY1(I,U(I),K)
50 WRITE(6,200) (Y(M),M=1,MN)
100 FORMAT(2I5)
101 FORMAT(I5)
102 FORMAT(2F10.2)
103 FORMAT(2I5,F10.2,I5)
200 FORMAT(10F8.2)
      END

      SUBROUTINE NORVEC(MN,E)
      DIMENSION E(1)
      DO 5 I=1,MN
      R1=RNRF(-1)
      R2=RNRF(-1)
5 E(I)=(-2.0*ALOG(R1))*+.5*COS(6.2831*R2)
      RETURN
      END

      FUNCTION TAREX(VAL,ARG,DUMMY,K)
C THIS FUNCTION SUBPROGRAM IS FROM LLEWELLYN(1965).
      DIMENSION VAL(1),ARG(1)
      DO 1 J=2,K
      IF(DUMMY.GT.ARG(J)) GO TO 1
1 TABEX=(DUMMY-ARG(J-1))*(VAL(J)-VAL(J-1))/
      1*(ARG(J)-ARG(J-1))+VAL(J-1)
      RETURN
1 CONTINUE
      J=K
      GO TO 2
      END

```

Figure A.9 A Generalized Multivariate Process Generator

Figure A.9 (cont'd.)

```

FUNCTION TABEXI(I, DUMMY, K)
COMMON /BLOCK2/ ARG(50, 100), VAL(50, 100)
DO 1 J=2, K
IF (DUMMY GT VAL(I, J)) GO TO 1
2 TABEXI=(DUMMY-VAL(I, J-1))*(ARG(I, J)-ARG(I, J-1))/
1 (VAL(I, J)-VAL(I, J-1))+ARG(I, J-1)
RETURN
1 CONTINUE
J=K
GO TO 2
END

SUBROUTINE COEF
C THIS SUBROUTINE CONSTRUCTS THE LOWER TRIANGULAR MATRIX C.
C IT WAS WRITTEN BY R NEAL PETERSON
COMMON /BLOCK1/ C(50, 50), COR(50, 50), PAR(50, 50), MN
DO 10 I=1, MN
DO 10 J=1, MN
10 C(I, J)=0
DO 20 I=1, MN
PAR(I, 1)=COR(I, 1)
20 IF (MN LT 2) GO TO 70
DO 60 J=2, MN
JM1=J-1
SUM=0
DO 30 K=1, JM1
SUM=SUM+PAR(J, K)*1.7
30 PAR(J, J)=SQRT(1 -SUM)
IF (J EQ MN) GO TO 70
JP1=J+1
DO 50 I=JP1, MN
PAR(I, J)=0
IF (PAR(J, J) EQ 0.) GO TO 50
SUM=0
DO 40 K=1, JM1
SUM=SUM+PAR(I, K)*PAR(J, K)
40 PAR(I, J)=(COR(I, J)-SUM)/PAR(J, J)
50 CONTINUE
60 CONTINUE
70 DO 80 I=1, MN
DO 80 J=1, 1
80 C(I, J)=PAR(I, J)
RETURN
END

```

marginal distribution need to be specified rather than specific points in the cumulative distribution function. In this study, for example, it was often necessary to generate sample vectors from a multivariate beta distribution. Accordingly, a special multivariate beta process generator was developed. A listing of program MVBETA, which implements this procedure, is given in Figure A.10.

Program MVBETA is identical to MVGEN in most respects. It differs only in that parameter values for each marginal beta distribution--BL, BU, K1, and K2--are supplied by the user instead of actual points on the cumulative distribution function.¹ The program then uses numerical integration techniques to construct the cumulative distribution function of each marginal. The other inputs to MVBETA are the same as those required for MVGEN.

A.5 An Empirical Test of the Generalized Multivariate Process Generator

The procedure described in the preceding section is based on the premise that correlations between random variables are, to a large extent, preserved as the variables undergo similar transformations. A simple experiment was designed and conducted to test the validity of this premise. Program MVBETA was used to generate 1000 sample vectors from a multivariate beta distribution with marginal distribution parameters and correlation matrix equal to those defined in Tables A.1 and A.2. It should be noted that each marginal distribution is skewed. The first three are negatively skewed, while the fourth is positively

¹These parameters are defined as in the preceding discussion of the univariate beta process generator.

```

PROGRAM MVNMTA(INPUT,OUTPUT,TAPE5=INPUT,TAPE6=OUTPUT)
COMMON /M00K1/ C(50,50),COR(50,50),PAR(50,50),MN
COMMON /M00K2/ ARG(50,100),VAL(50,100)
DIMENSION RV(50),U(50),E(50),VALN(41),ARGN(41)
DIMENSION Y(50)
C THE NEXT TWO DATA STATEMENTS ASSIGN VALUES TO VARIABLES
C USED TO CONSTRUCT A TABLE LOOK UP REPRESENTATION OF THE
C INVERSE OF THE CUMULATIVE DISTRIBUTION FUNCTION OF A
C STANDARD NORMAL RANDOM VARIABLE.
DATA KX/41/
DATA ARGN/3 5,-1.96,-1.645,-1.439,-1.281,-1.150,-1.037,-.925,
1-.841,-.750,-.674,-.598,-.524,-.454,-.386,-.312,-.253,-.189,
2-.125,-.064,0,.054,.136,.197,.253,.312,.386,.454,.524,.598,
3-.674,-.750,-.841,-.925,-1.037,-1.150,-1.281,-1.439,-1.645,-1.960,3.5/
C READ THE NUMBER OF SAMPLE VECTORS TO BE GENERATED, ND, AND
C THE NUMBER OF ELEMENTS IN EACH VECTOR, MN.
READ(5,100) ND,MN
C READ THE NUMBER OF POINTS TO BE DEFINED ON THE CUMULATIVE
C DISTRIBUTION OF EACH MARGINAL, K.
READ(5,101) K
DARC=1.0/VAL(K-1)
DB=.10/DARC
DO 5 I=1,MN
C READ PARAMETERS FOR EACH MARGINAL DISTRIBUTION.
READ(5,100) RL,BU,K1,K2
C USE EULER INTEGRATION TO INTEGRATE THE DENSITY FUNCTION OF
C EACH MARGINAL IN THIS MANNER THE CUMULATIVE DISTRIBUTION
C FUNCTION OF EACH MARGINAL IS CONSTRUCTED
ARG(I,1)=RL
VAL(I,1)=0.0
RANGE=BU-RL
CA=GAMMA(K1)
CB=GAMMA(K2)
KA=K1-1
KB=K2-1
KD=K1+K2
CD=GAMMA(KD)
CON=CD/(CA*CB)
CON=CON*DB
B=0.0
P=0.0
DO 5 J=2,K
DO 1 M=1,10
P=P+CON*(F**KA)*((1.0-B)**KB)
1 B=B+DB
ARG(I,J)=ARG(I,J-1)+DARC*RANGE
VAL(I,J)=P
IF(VAL(I,J) GT 0.999999) VAL(I,J)=0.999999
IF(J LT K) GO TO 5
ARG(I,J)=BU
VAL(I,J)=1.0
5 CONTINUE
C SET ALL OFF DIAGONAL ELEMENTS OF THE CORRELATION MATRIX TO ZERO.
DO 10 I=1,MN
DO 10 J=1,MN
COR(I,J)=0.0
IF(I EQ J) COR(I,J)=1.0
10 CONTINUE
C READ ALL NON-ZERO CORRELATION COEFFICIENTS. SET IND EQUAL
C TO A NON-ZERO VALUE WHEN THE LAST CORRELATION COEFFICIENT IS
C TO BE READ.
15 READ(5,100) I,J,COR(I,J),IND
COR(J,I)=COR(I,J)
IF(IND EQ 0) GO TO 15
C CONSTRUCT LOWER TRIANGULAR MATRIX C.
CALL CORR
VALN(1)=0.0
DO 20 I=2,MN
J=I-1
20 VALN(I)=VALN(J)+.025
C BEGIN TO GENERATE RANDOM VARIATES.
DO 50 M=1,MN
C CALL NORVEC, WHICH RETURNS A VECTOR OF INDEPENDENTLY DISTRIBUTED
C STANDARD NORMAL RANDOM VARIATES.
CALL NORVEC(COR)
C CONSTRUCT SAMPLE VECTOR FROM A MULTIVARIATE NORMAL DISTRIBUTION.
DO 30 I=1,MN
RV(I)=0.0
DO 25 J=1,MN
25 RV(I)=RV(I)+C(I,J)*E(J)
30 CONTINUE
C TRANSFORM THE VECTOR OF MULTIVARIATE NORMAL RANDOM VARIABLES
C INTO A VECTOR OF UNIFORMLY DISTRIBUTED RANDOM VARIABLES.
DO 35 I=1,MN
35 U(I)=1+EXP(VALN(ARGN,I)*E(I),KN)
C TRANSFORM THE UNIFORMLY DISTRIBUTED RANDOM VARIABLES INTO
C RANDOM VARIABLES FROM THE SPECIFIED BETA DISTRIBUTIONS.
DO 40 I=1,MN
40 Y(I)=TADFX(1,U(I),K)
50 WRITE(6,200) (Y(M),M=1,MN)
100 FORMAT(215)
101 FORMAT(15)
102 FORMAT(15,2,215)
103 FORMAT(215,10,2,15)
200 FORMAT(10F5,2)
END

```

Figure A.10 A Multivariate Beta Process Generator

Figure A.10 (cont'd.)

```

SUBROUTINE NORVEC(MN,E)
  DIMENSION E(1)
  DO 5 I=1,MN
    R1=RANF(-1)
    R2=RANF(-1)
5    E(I)=(-2.0*ALOG(R1))*5*COS(6.2831*R2)
  RETURN
END

FUNCTION TABEX(VAL,ARG,DUMMY,K)
C THIS FUNCTION SUBPROGRAM IS FROM LLEWELLYN(1965).
  DIMENSION VAL(1),ARG(1)
  DO 1 J=2,K
    IF(DUMMY GT ARG(J)) GO TO 1
2    TABEX=(DUMMY-ARG(J-1))*(VAL(J)-VAL(J-1))/
1    (ARG(J)-ARG(J-1))+VAL(J-1)
  RETURN
1 CONTINUE
  J=K
  GO TO 2
END

FUNCTION TABEXI(I,DUMMY,K)
  COMMON /BLOCK2/ ARG(50,100),VAL(50,100)
  DO 1 J=2,K
    IF(DUMMY GT VAL(I,J)) GO TO 1
2    TABEXI=(DUMMY-VAL(I,J-1))*(ARG(I,J)-ARG(I,J-1))/
1    (VAL(I,J)-VAL(I,J-1))+ARG(I,J-1)
  RETURN
1 CONTINUE
  J=K
  GO TO 2
END

FUNCTION GAMMA(K)
  KK=K-1
  A=1.0
  IF(KK LE 1) GO TO 10
  DO 5 I=2,KK
    A=A*FLOAT(I)
5  GAMMA=A
10 RETURN
END

SUBROUTINE COEF
C THIS SUBROUTINE CONSTRUCTS THE LOWER TRIANGULAR MATRIX C.
C IT WAS WRITTEN BY R. NEAL PETERSON
  COMMON /BLOCK1/ C(50,50),COR(50,50),PAR(50,50),MN
  DO 10 I=1,MN
    DO 10 J=1,MN
10    C(I,J)=0
  DO 20 I=1,MN
    PAR(I,1)=COR(I,1)
    IF(MN LT 2) GO TO 70
    DO 60 J=2,MN
      JM1=J-1
      SUM=0.
30    DO 30 K=1,JM1
      SUM=SUM+PAR(J,K)**2
      PAR(J,J)=SQRT(1.-SUM)
      IF(J EQ MN) GO TO 70
      JP1=J+1
      DO 50 I=JP1,MN
        PAR(I,J)=0.
        IF(PAR(J,J) EQ 0.) GO TO 50
      SUM=0.
40    DO 40 K=1,JM1
      SUM=SUM+PAR(I,K)*PAR(J,K)
      PAR(I,J)=(COR(I,J)-SUM)/PAR(J,J)
50    CONTINUE
60    CONTINUE
70    DO 80 I=1,MN
      DO 80 J=1,I
80    C(I,J)=PAR(I,J)
  RETURN
END

```

Table A.1 Parameters of Marginal Beta Distributions^a

| Distribution | K1 | K2 | BL | BU | μ | σ |
|--------------|----|----|------|-------|-------|----------|
| 1 | 2 | 1 | 22.0 | 31.0 | 28.0 | 2.12 |
| 2 | 3 | 2 | 70.0 | 110.0 | 91.0 | 8.00 |
| 3 | 3 | 2 | 4.50 | 9.00 | 7.20 | .90 |
| 4 | 1 | 4 | 1.75 | 3.20 | 2.04 | .24 |

^aDistributions 1 and 2 can be considered to be national average yields for soybean and corn respectively, while distributions 3 and 4 can be considered to be national average prices for the two respective crops.

Table A.2 Specified Correlation Matrix

| Distribution | 1 | 2 | 3 | 4 |
|--------------|------|------|------|------|
| 1 | 1 | .75 | -.30 | -.20 |
| 2 | .75 | 1 | -.20 | -.30 |
| 3 | -.30 | -.20 | 1 | .45 |
| 4 | -.20 | -.30 | .45 | 1 |

skewed. Program MVNOR was then used to generate 1000 sample vectors from a multivariate normal distribution having marginal distributions with means and variances identical to those specified for the marginal beta distributions in Table A.1 and a correlation matrix identical to that of the multivariate beta distribution. The sample correlation matrix for each set of observations was then calculated and compared to that which was specified. The procedure for generating sample observations from the multivariate normal distribution was derived analytically, and, in the limit, the sample correlation matrix for data generated with MVNOR should converge to the correlation matrix actually specified. Therefore, the sample properties of the observations generated with MVNOR provide a standard against which the performance of MVBETA can be judged.

The sample correlation matrix for the data generated with MVNOR is shown in Table A.3. For the most part, it conforms closely to that specified. The mean absolute deviation for off diagonal elements is only .016. The sample correlation matrix for the data generated with MVBETA is shown in Table A.4. It, too, conforms closely to that specified, and the mean absolute deviation for off diagonal elements of .018 is only slightly larger than that for the data generated with MVNOR.

These results indicate that, at least in this particular case, correlations between random variables are preserved to a large extent as the variables undergo the series of transformations required to make their marginal distributions take a specified form. Clearly additional tests should be performed, but all experience to date with the generalized multivariate process generator has suggested the results presented here.

Table A.3 Sample Correlation Matrix for
Multivariate Normal Data

| Distribution | Distribution | | | |
|--------------|--------------|-------|-------|-------|
| | 1 | 2 | 3 | 4 |
| 1 | 1 | .765 | -.320 | -.195 |
| 2 | .765 | 1 | -.180 | -.295 |
| 3 | -.320 | -.180 | 1 | .417 |
| 4 | -.195 | -.295 | .417 | 1 |

Table A.4 Sample Correlation Matrix for
Multivariate Beta Data

| Distribution | Distribution | | | |
|--------------|--------------|-------|-------|-------|
| | 1 | 2 | 3 | 4 |
| 1 | 1 | .753 | -.313 | -.192 |
| 2 | .753 | 1 | -.174 | -.291 |
| 3 | -.313 | -.174 | 1 | .404 |
| 4 | -.192 | -.291 | .404 | 1 |

Given the degree of imprecision with which correlation coefficients are generally specified, the inaccuracies associated with the use of this approach can hardly be considered to be significant.

A.6 Conclusions

The multivariate process generator presented here is a reliable, easily implemented procedure which greatly enhances the value of Monte Carlo simulation techniques in the analysis of decisions made under uncertainty. It facilitates the representation of stochastic processes which cannot accurately be described by a multivariate normal distribution and, when used in conjunction with other simulation techniques, permits the incorporation of a greater degree of realism into the analysis of complex problems. As such, its contribution to the set of techniques developed in this study is an important one.

APPENDIX B

IMPLEMENTATION OF THE INTERVAL APPROACH TO THE MEASUREMENT OF DECISION MAKER PREFERENCES

B.1 Introduction

Implementation of the interval approach to the measurement of decision maker preferences was discussed in Section 4.6 of Chapter IV. In this appendix a more technical explanation of its implementation is presented, with particular attention being given to the description of several computer programs which greatly facilitate the process. More specifically, six steps in the implementation of this preference measurement technique are identified and discussed: the specification of a measurement scale, the generation of sample distributions, the identification of a boundary interval for each pair of distributions, the construction of the questionnaire, administration of the questionnaire, and the use of interval preference measurements to order alternative choices.

B.2 Specification of a Measurement Scale

Implementation of the interval approach begins with the specification of a measurement scale--a set of reference levels of absolute risk aversion which serve as the basis for preference measurements. Because this scale determines the degree of precision with which preference measurements can be made, careful attention should be given to its specification. Two related questions which must be considered at

this point are: how many reference levels to include in the measurement scale and where to locate these levels on the scale.

The number of reference levels on the measurement scale depends to a large extent on the number of choices the decision maker will be asked to make in measuring absolute risk aversion in the neighborhood of any particular system output level. There is a direct relationship between the number of questions to be asked and the minimum number of reference levels on the measurement scale. If the decision maker is required to make only one choice, only one boundary interval need be specified, and only two reference levels are required to define that interval. If a second choice is to be required, two additional boundary intervals must be specified, since the interval which is the focus of the second question will depend on the decision maker's response to the first. Four reference levels are required to define the total of three boundary intervals called for in this case. Similarly, seven distinct boundary intervals defined by eight reference levels are required for the construction of a three question sequence, since a separate boundary interval is the focus of the choice made at each node of the hierarchy of questions defined in Figure B.1. It follows by similar reasoning that if N Choices are to be made in measuring absolute risk aversion in the neighborhood of a particular system output level, the measurement scale should be comprised of at least 2^N reference levels.

Once the number of reference levels to be specified has been determined, their location in risk aversion space must be established. Reference levels on the measurement scale need not be placed at regular intervals. As was noted in Chapter IV, it is usually advisable to concentrate reference levels in that region of absolute risk aversion

Question 1

A vs B

A

Question 2

C vs D

C

Question 3

G vs H or

I vs J

or

K vs L

or M vs N

B

E vs F

E

F

Figure B.1 Sequence of Choices for a Measurement of Absolute Risk Aversion Based on Three Questions

space where the decision maker's actual level is expected to fall or in the regions where relatively small changes in absolute risk aversion have the greatest impact on preference orderings. Experience to date indicates that most of the detail on the measurement scale should be concentrated in the risk aversion interval between $-.0001$ and $.0010$. Actual measurements for a variety of decision makers have tended to fall most frequently within this interval, and tests on several empirical decision problems have indicated that choices are most strongly affected by changes in absolute risk aversion within this range. A suggested set of sixteen reference levels is given in Figure B.2. These define fifteen boundary intervals upon which choices in a four question sequence could be focused. A measurement scale for a three question sequence could be constructed by using every other reference level, and that for a two question sequence could be constructed by using every fourth reference level.

B.3 The Generation of Sample Distributions

Once a measurement scale has been specified, the sample probability distributions which are the basis for the choices used to reveal the decision maker's preferences must be generated. Program NORGEN, which is listed in Figure B.3, is used to construct these distributions, each of which is actually a set of sample observations drawn from a normal distribution with specified mean and standard deviation.¹ These sample

1

A normal distribution is used because it is convenient. Any other underlying distribution can also be the basis for the generation of sample distributions.

244

.0100

.0050

.0025

.0015

.0010

.0008

.0006

.0004

.0003

.0002

.0001

0

-.0001

-.00025

-.0005

-.0010

Figure B.2 A Suggested Absolute Risk
Aversion Measurement Scale

```

      PROGRAM NORGEN(INPUT,OUTPUT,TAPES=INPUT,TAPE4)
C READ NUMBER OF DISTRIBUTIONS TO BE GENERATED, NE, NUMBER
C OF SAMPLE OBSERVATIONS IN EACH DISTRIBUTION, ND, MEAN OF
C THE UNDERLYING DISTRIBUTION, YMEAN, STANDARD DEVIATION OF
C THE UNDERLYING DISTRIBUTION, STD, AND THE ROUNDING FACTOR,
C IROUND.
      READ(5,200) NE,ND,YMEAN,STD,IROUND
      ISTD=STD/FLOAT(IROUND) + .99999
C GENERATE SAMPLE DISTRIBUTIONS.
      DO 20 I=1,NE
        PRINT 100,I
        WRITE(4,100) I
        DO 10 J=1,ND
          R1=PRNF(-1)
          R2=PRNF(-1)
          X=(-2.0*ALOG(R1))*+.5*CCS(6.2831*R2)
          Y=FLOAT(ISTD)*X
          Y=FLOAT(J*IROUND)+YMEAN
          WRITE(4,101) Y
          PRINT 101,Y
        10 CONTINUE
      20 CONTINUE
      100 FORMAT(* DIST*,I2)
      101 FORMAT(F10.2)
      200 FORMAT(2I5,2F10.2,I5)
      END

```

Figure B.3 A Listing of Program NORGEN

observations are considered to be possible system output levels, and each is said to have an equal probability of occurrence.

The user of NORGEN must specify values for five parameters: NE, ND, YMEAN, STD, and IROUND. NE is the number of sample distributions to be generated, and ND is the number of sample observations defining each distribution. Recommended values for NE and ND are 40 and 6 respectively. The generation of forty sample distributions almost guarantees that at least one pair of distributions will have its boundary interval at any specified level. The use of six-element distributions is justified by the ease of explaining the probability associated with each element and by the fact that this number of elements is sufficient to allow for considerable complexity in each distribution. YMEAN and STD are the mean and standard deviation of the underlying normal distribution from which the sample elements of each distribution are drawn. In practice, YMEAN is usually set equal to 0.0, which implies that the expected value of the mean of each sample distribution is also 0.0. That expected value can be shifted to any level, y^* , however, by simply adding y^* to each element of a distribution. The appropriate value for STD depends on the characteristics of the decision situation being analyzed. If STD is assigned too high a value, the dispersion of the sample distributions will be great and the assumption of constant absolute risk aversion over the range of system output levels on which they are defined may be difficult to justify. If STD is assigned too low a value, on the other hand, the points defining each distribution will be so highly concentrated around a single system output level that choices between distribution will be difficult to make. Experience to date indicates

that a value of STD between one and five percent of the entire relevant range of system output levels is appropriate. Finally, it is often desirable to round the system output levels defining each sample observation to the nearest 10, 50, or 100 units. This can be accomplished by specifying a value for IROUND. If IROUND is set equal to 50, which is the recommended value, all system output levels are rounded down to the nearest 50 units. A sample output from program NORGEN is given in Figure B.4.

B.4 Identification of Boundary Intervals

After a measurement scale has been specified and sample distributions have been generated, the boundary interval for each pair of distributions must be identified. The boundary interval for two distributions, (λ_1, λ_2) , is an interval in risk aversion space such that decision makers whose absolute risk aversion functions lie everywhere below λ_1 unanimously prefer one distribution, while those whose absolute risk aversion functions lie everywhere above λ_2 unanimously prefer the other. Clearly a boundary interval is not unique. If (λ_1, λ_2) is a boundary interval for two distributions, for example, and if $\lambda_3 < \lambda_1$ and $\lambda_4 > \lambda_2$, then (λ_3, λ_4) is also a boundary interval for these two distributions. In measuring preferences, however, it is desirable to specify boundary intervals which are as narrow as possible.

The absolute risk aversion reference levels which define the measurement scale constitute the set of potential endpoints for boundary intervals. If a measurement scale is comprised of four reference levels, $-.0010$, 0 , $.0005$, and $.0010$, a total of six boundary intervals can be constructed: $(-.0010, 0)$, $(0, .0005)$, $(.0005, .0010)$, $(-.0010, .0005)$,

```

DIST 1
-450.00
-100.00
600.00
-250.00
-300.00
450.00
DIST 2
400.00
-550.00
1100.00
0.00
0.00
650.00
DIST 3
-300.00
-550.00
300.00
-650.00
150.00
-450.00
DIST 4
200.00
100.00
-200.00
-450.00
50.00
-300.00
DIST 5
-600.00
50.00
150.00
-100.00
-100.00
-150.00
DIST 6
550.00
-350.00
-450.00
-650.00
600.00
-50.00
DIST 7
450.00
1100.00
-450.00
400.00
-1000.00
-150.00
DIST 8
200.00
-50.00
50.00
150.00
-950.00
0.00
DIST 9
-400.00
-350.00
-750.00
50.00
1000.00
150.00
DIST 10
-400.00
350.00
450.00
150.00
300.00
-450.00
DIST 11
-250.00
-450.00
-150.00
250.00
300.00
300.00
DIST 12
-350.00
500.00
100.00
-150.00
-250.00
-150.00
DIST 13
-300.00
-350.00
-150.00
400.00
800.00
-50.00
DIST 14
-550.00
-700.00
500.00
-50.00
450.00
-100.00
DIST 15
-250.00
-700.00
-400.00
-300.00
-700.00
-400.00
DIST 16
-450.00
-950.00
-550.00
-50.00
-100.00
50.00
DIST 17
700.00
-850.00
-1000.00
0.00
1100.00
-950.00
DIST 18
150.00
0.00
-600.00
-950.00
-250.00
100.00
DIST 19
-100.00
1000.00
-750.00
150.00
250.00
100.00
DIST 20
600.00
400.00
-700.00
-750.00
300.00
-400.00

```

Figure B.4 Sample Output from Program NORGEN

(0, .0010), and (-.0010, .0010). Because relatively narrow boundary intervals are sought, however, only the first three--those defined by adjacent reference levels--are of interest. For any pair of sample distributions it is necessary to determine which, if any, of these three intervals can be said to be a boundary interval. Given the definition of a boundary interval, this requires the identification of the highest reference level, λ_1 , such that all decision makers less risk averse than λ_1 prefer one distribution and the lowest reference level, λ_2 , such that all decision makers more risk averse than λ_2 prefer the other distribution.

Program INTID, which is listed in Figure B.5, is used to accomplish this task. Given a set of absolute risk aversion reference levels and a set of sample distributions, it identifies the narrowest boundary interval for each pair of distributions. It does this by applying stochastic dominance criteria developed by Meyer (1977b) in "Second Degree Stochastic Dominance with Respect to a Function." Subroutine SDLB of INTID orders distributions for classes of decision makers whose absolute risk aversion functions are bounded only from below by applying the following criterion: cumulative distribution function $F(y)$ is unanimously preferred to cumulative distribution function $G(y)$ by all decision makers more risk averse than $k(y)$ if and only if:

$$\int_{-\infty}^y [G(x) - F(x)] dk(x) \geq 0 \quad \forall y \quad (B.1)$$

This subroutine, then, is used to identify the upper bound of the boundary interval. Subroutine SDUB, on the other hand, orders distributions

¹This criterion is based on Meyer's (1977b, p. 479) Definition 4.

```

      PROGRAM INTID(INPUT,OUTPUT,TAPE7=INPUT,TAPE6=OUTPUT,TAPE4)
      COMMON /BLOCK1/ ND,C(21),CP(21),R(40,10),RA(50)
      DIMENSION NAME(40),T(10)
      C READ NUMBER OF DISTRIBUTIONS TO BE CONSIDERED, NE, NUMBER
      C OF SAMPLE OBSERVATIONS DEFINING EACH DISTRIBUTION, ND, AND
      C NUMBER OF LEVELS ON MEASUREMENT SCALE.
      READ(5,250) NE,ND,NG
      250 FORMAT(3I5)
      NG=NG-1
      WRITE(4,300)
      300 FORMAT(1H,3X,NAME,40X,DATA POINTS,45X,MEAN,9X,STD)
      C READ DATA ON DISTRIBUTIONS TO BE CONSIDERED.
      DO 100 I=1,NE
      READ(4,400) NAME(I)
      400 FORMAT(A10)
      DO 500 J=1,ND
      READ(4,500) R(I,MMM)
      500 CONTINUE
      5001 FORMAT(F10,2)
      C COMPUTE MEAN AND STANDARD DEVIATION OF EACH SAMPLE DISTRIBUTION.
      VAR=0
      AMEAN=0
      PAD=0
      DO 101 J=1,ND
      AMEAN=AMEAN+(1.0/PAD)*R(I,J)
      VAR=VAR+(1.0/PAD)*R(I,J)*R(I,J)
      T(J)=R(I,J)
      101 CONTINUE
      VAR=VAR-A*AMEAN
      STD=SQRT(VAR)
      400 FORMAT(1H,410,10(1X,F8.2),2(4X,F8.2))
      C SORT DISTRIBUTIONS.
      CALL BORFER(T)
      DO 102 J=1,ND
      R(I,J)=T(J)
      102 CONTINUE
      100 CONTINUE
      WRITE(4,305)
      305 FORMAT(1H1)
      C READ ABSOLUTE RISK AVERSION LEVELS WHICH DEFINE MEASUREMENT
      C SCALE.
      READ(5,201) (RA(I),I=1,NG)
      201 FORMAT(5F10,6)
      WRITE(6,251)
      251 FORMAT(5X,ABSOLUTE RISK AVERSION LEVELS DEFINING MEASUREMENT SCA
      1LE)
      253 WRITE(6,252) RA(M)
      252 FORMAT(1H,20X,F10,6)
      M=M-1
      IF(M.GT.0) GO TO 253
      601 FORMAT(1H1,20X,SPECIFICATION OF BOUNDARY INTERVALS)
      C IDENTIFY BOUNDARY INTERVAL FOR EACH PAIR OF DISTRIBUTIONS.
      DO 105 I=1,NE
      DO 110 J=2,NE
      IF(J.LE.I) GO TO 110
      CALL CUMCAL(J,I)
      DO 110 K=1,NG
      L=K+1
      CALL SPUR(K,INDL)
      CALL SCLP(L,INDU)
      IF(INDL.GE.0.CP,INDU,FC,0) GO TO 110
      IF(INDL.GE.0,INDU) GO TO 110
      IF(INDL.GE.-1.AND.INDU,FC,1) GO TO 105
      WRITE(6,400) RA(K),RA(L),NAME(I),RA(K),NAME(J),RA(L)
      GO TO 110
      105 WRITE(6,400) RA(K),RA(L),NAME(J),RA(K),NAME(I),RA(L)
      400 FORMAT(1H,410,5,10,5,2X,A10,2X,PREFERRED BELOW,5,F10,5,2X,
      1A10,2X,PREFERRED ABOVE,5,F10,5)
      GO TO 110
      110 CONTINUE
      115 CONTINUE
      120 CONTINUE
      END

      FUNCTION U(X,JJ)
      COMMON /BLOCK1/ ND,C(21),CP(21),R(40,10),RA(50)
      DE=RA(JJ)*X
      IF(RA(JJ).92.91.90)
      90 U=-EXP(0)
      GO TO 93
      91 U=X
      GO TO 93
      92 U=EXP(0)
      93 RETURN
      END

```

Figure B.5 A Listing of Program INTID

Figure B.5 (cont'd.)

```

SUBROUTINE SOLP(K,IND)
COMMON /FLOCK1/ ND,C(21),CP(21),R(47,10),RA(50)
DIF=C(0)
ND2=ND*2
IFLAG=0
IND=0
DO 5 M=1,ND2
DIF=IF-C(M)*(U(CP(M+1),K)-U(CP(M),K))
IF(ABS(DIF) .GT. 10,5,4)
4 TELCC=1
5 CONTINUE
IF(FLAG.EQ.1) IND=1
RETURN
10 IF(FLAG.EQ.1) GO TO 20
DIF=C(0)
DO 15 M=1,ND2
DIF=IF-C(M)*(U(CP(M+1),K)-U(CP(M),K))
IF(ABS(DIF) .GT. 20,15,15)
15 CONTINUE
IND=-1
20 RETURN
END

```

```

SUBROUTINE SDUP(K,IND)
COMMON /FLOCK1/ ND,C(21),CP(21),R(47,10),RA(50)
DIF=C(0)
ND2=ND*2
NLEN=0
IFLAG=0
IND=0
DO 5 NL=1,21 GO TO 5
DIF=IF-C(NL)*(U(CP(NL+1),K)-U(CP(NL),K))
NLEN=NLEN+1
IF(ABS(DIF) .GT. 4,5,10)
4 IFLAG=1
5 CONTINUE
IF(FLAG.EQ.1) IND=-1
RETURN
10 IF(FLAG.EQ.1) GO TO 20
DIF=C(0)
NLEN=0
DO 15 NL=1,21 GO TO 15
DIF=IF-C(NL)*(U(CP(NL+1),K)-U(CP(NL),K))
NLEN=NLEN+1
IF(ABS(DIF) .GT. 15,15,20)
15 CONTINUE
IND=1
20 RETURN
END

```

```

SUBROUTINE RCRDEF(T)
COMMON /FLOCK1/ ND,C(21),CP(21),R(47,10),RA(50)
DIMENSION T(1)
ND1=ND-1
DO 401 I=1,ND1
AM=T(I)
M=1
II=I+1
DO 401 J=II,ND
IF(AM.LE.T(J)) GO TO 401
M=J
401 CONTINUE
T(I)=T(1)
T(I)=AM
402 CONTINUE
RETURN
END

```

Figure B.5 (cont'd.)

```

SUBROUTINE CUMCAL(M1,M2)
COMMON /FLOCK1/ ND,C(21),CP(21),P(43,10),RA(50)
K1=1
K2=1
C(1)=0.0
CP(1)=0.0
ND=ND
ND2=ND*2+1
DO 500 I=2,ND2
IF (K1.GT.ND.AND.K2.GT.ND) GO TO 550
IF (K1.GT.ND) GO TO 553
IF (K2.GT.ND) GO TO 551
IF (R(M1,K1)-R(M2,K2)) 551,552,553
551 C(I)=C(I-1)+(1.0/PA0)
CP(I)=C(I,K1)
K1=K1+1
GO TO 500
552 C(I)=C(I-1)
CP(I)=R(M1,K1)
K1=K1+1
K2=K2+1
GO TO 500
553 C(I)=C(I-1)-(1.0/5ND)
CP(I)=R(M2,K2)
K2=K2+1
GO TO 500
550 C(I)=0.0
CP(I)=0.0
500 CONTINUE
RETURN
END

```

for classes of decision makers defined only by an upper bound in absolute risk aversion using the criterion: cumulative distribution function $G(y)$ is preferred to cumulative distribution function $F(y)$ by all decision makers less risk averse than $k(y)$ if and only if:

$$\int_y^{\infty} [G(x) - F(x)] dk(x) \leq 0 \quad \forall y^1 \quad \text{B.2}$$

This subroutine is used to identify the lower bound of the boundary interval.

In determining the boundary interval for a particular pair of distributions, program INTID tests each interval on the measurement scale until a boundary interval defined by two adjacent reference levels is identified. Given the reference levels--.0010, 0, .0005, and .0010, for example, the interval $(-.0010, 0)$ is considered first. If subroutine SDUB indicates unanimous preference for one distribution at absolute risk aversion levels below 0.0010 and subroutine SDLB indicates unanimous preference for the other at absolute risk aversion levels above 0, $(0.0010, 0)$ is a boundary interval. If this criterion is not met, the interval $(0, .0005)$ is evaluated in the same manner. The program steps up the measurement scale in this way until a boundary interval is identified or until all possible intervals have been examined.

Several parameter values must be specified by the user of INTID. NE and ND are defined exactly as in program NORGEN; they indicate the number of distributions to be considered and the number of elements defining each distribution. Maximum values of NE and ND are 40 and 10

¹This criterion is based on Meyer's (1977b, p. 482) Theorem 5.

respectively. NG is the number of reference levels on the measurement grid. Its maximum value is 64, the number of reference levels required for the specification of a six question sequence. Two other types of inputs must be supplied by the user of INTID. First, the output of program NORGEN--the NE distributions, each having ND elements--must be read into INTID.¹ These data are stored in two arrays: NAME, an array of distribution names, and R, an array of sample points. Second, the reference levels defining the measurement scale must be specified and read into array RA. A sample output from program INTID identifying boundary intervals for the distributions given in Figure B.4 based on the measurement scale defined in Figure B.6 is given in Figure B.7.

B.5 Construction of the Questionnaire

At least one pair of distributions for which the boundary interval lies between any two adjacent reference levels on the measurement scale should be identified by program INTID. Once this has been done, a hierarchy of questions can be established, with each question focusing on a different boundary interval. The hierarchy of questions associated with the measurement scale defined in Figure B.7 is given in Figure B.8.

In general the first question of such a hierarchy should focus on the boundary interval at the center of the measurement scale. That in Figure B.8, for example, focuses on the boundary interval (.0001, .0002), which is defined by the fourth and fifth reference levels of the eight-level measurement scale. The two questions at the second level focus

¹Programs NORGEN and INTID are written so that the output of NORGEN can be catalogued as a permanent file and read into INTID from that file.

.0050
.0010
.0006
.0003
.0001
0
-.0001
-.0005

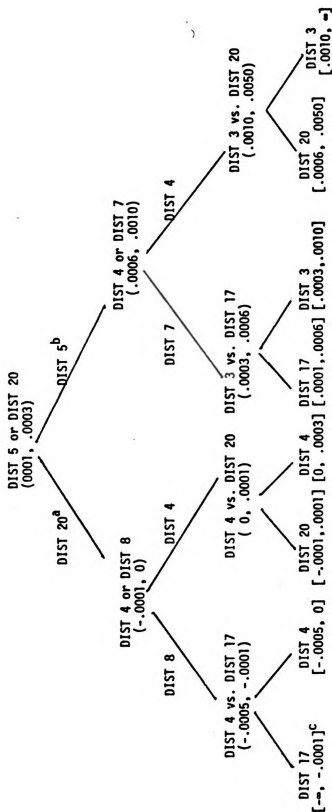
Figure B.6 An Eight Element Measurement Scale

| SPECIFICATION OF BOUNDARY INTERVALS | | | | | |
|-------------------------------------|---------|---------|-----------------|---------|---------|
| .00100 TO | .00500 | DIST 2 | PREFERRED BELOW | .00100 | DIST 1 |
| -.00050 TO | -.00010 | DIST 19 | PREFERRED BELOW | -.00050 | DIST 1 |
| .00100 TO | .00500 | DIST 2 | PREFERRED BELOW | .00100 | DIST 11 |
| .00100 TO | .00500 | DIST 2 | PREFERRED BELOW | .00100 | DIST 12 |
| .00100 TO | .00500 | DIST 2 | PREFERRED BELOW | .00100 | DIST 13 |
| .00100 TO | .00500 | DIST 7 | PREFERRED BELOW | .00100 | DIST 3 |
| .00100 TO | .00500 | DIST 9 | PREFERRED BELOW | .00100 | DIST 3 |
| .00100 TO | .00500 | DIST 16 | PREFERRED BELOW | .00100 | DIST 3 |
| .00030 TO | .00260 | DIST 17 | PREFERRED BELOW | .00030 | DIST 3 |
| .00100 TO | .00500 | DIST 20 | PREFERRED BELOW | .00100 | DIST 3 |
| .00030 TO | .00060 | DIST 6 | PREFERRED BELOW | .00030 | DIST 4 |
| .00060 TO | .00100 | DIST 7 | PREFERRED BELOW | .00060 | DIST 4 |
| -.00010 TO | 0.00000 | DIST 8 | PREFERRED BELOW | -.00010 | DIST 4 |
| .00030 TO | .00060 | DIST 9 | PREFERRED BELOW | .00030 | DIST 4 |
| .00030 TO | .00260 | DIST 14 | PREFERRED BELOW | .00030 | DIST 4 |
| -.00050 TO | -.00010 | DIST 17 | PREFERRED BELOW | -.00050 | DIST 4 |
| .00100 TO | .00500 | DIST 19 | PREFERRED BELOW | .00100 | DIST 4 |
| 0.00000 TO | .00010 | DIST 20 | PREFERRED BELOW | 0.00000 | DIST 4 |
| .00060 TO | .00100 | DIST 6 | PREFERRED BELOW | .00060 | DIST 5 |
| .00060 TO | .00100 | DIST 7 | PREFERRED BELOW | .00060 | DIST 5 |
| .00030 TO | .00060 | DIST 8 | PREFERRED BELOW | .00030 | DIST 5 |
| .00060 TO | .00100 | DIST 9 | PREFERRED BELOW | .00060 | DIST 5 |
| .00060 TO | .00100 | DIST 14 | PREFERRED BELOW | .00060 | DIST 5 |
| -.00050 TO | -.00010 | DIST 17 | PREFERRED BELOW | -.00050 | DIST 5 |
| .00100 TO | .00500 | DIST 19 | PREFERRED BELOW | .00100 | DIST 5 |
| .00100 TO | .00500 | DIST 19 | PREFERRED BELOW | .00100 | DIST 5 |

Figure B.7 Sample Output for Program INTID

Figure B.7 (cont'd.)

| | | | | | | | |
|------------|---------|--------|-----------------|---------|--------|-----------------|---------|
| .00010 TC | .00030 | DIST20 | PREFERRED BELOW | .00010 | DIST 5 | PREFERRED ABOVE | .00030 |
| .00060 TC | .00100 | DIST 7 | PREFERRED BELOW | .00060 | DIST 6 | PREFERRED ABOVE | .00100 |
| -.00050 TC | -.00010 | DIST17 | PREFERRED BELOW | -.00050 | DIST 6 | PREFERRED ABOVE | -.00010 |
| .00100 TC | .00500 | DIST 7 | PREFERRED BELOW | .00100 | DIST 8 | PREFERRED ABOVE | .00500 |
| .00100 TC | .00500 | DIST 7 | PREFERRED BELOW | .00100 | DIST 9 | PREFERRED ABOVE | .00500 |
| -.00010 TC | 0.00000 | DIST 7 | PREFERRED BELOW | -.00010 | DIST10 | PREFERRED ABOVE | 0.00000 |
| -.00050 TC | -.00010 | DIST 7 | PREFERRED BELOW | -.00050 | DIST11 | PREFERRED ABOVE | -.00010 |
| .00010 TC | .00030 | DIST 7 | PREFERRED BELOW | .00010 | DIST12 | PREFERRED ABOVE | .00030 |
| .00100 TC | .00500 | DIST 7 | PREFERRED BELOW | .00100 | DIST14 | PREFERRED ABOVE | .00500 |
| .00100 TC | .00500 | DIST 7 | PREFERRED BELOW | .00100 | DIST15 | PREFERRED ABOVE | .00500 |
| .00100 TC | .00500 | DIST 7 | PREFERRED BELOW | .00100 | DIST16 | PREFERRED ABOVE | .00500 |
| .00100 TC | .00500 | DIST 7 | PREFERRED BELOW | .00100 | DIST18 | PREFERRED ABOVE | .00500 |
| .00100 TC | .00500 | DIST 7 | PREFERRED BELOW | .00100 | DIST20 | PREFERRED ABOVE | .00500 |
| .00100 TC | .00500 | DIST 8 | PREFERRED BELOW | .00100 | DIST15 | PREFERRED ABOVE | .00500 |
| -.00050 TC | -.00010 | DIST17 | PREFERRED BELOW | -.00050 | DIST 8 | PREFERRED ABOVE | -.00010 |
| .00060 TC | .00100 | DIST19 | PREFERRED BELOW | .00060 | DIST10 | PREFERRED ABOVE | .00100 |
| .00100 TC | .00500 | DIST11 | PREFERRED BELOW | .00100 | DIST12 | PREFERRED ABOVE | .00500 |
| .00060 TC | .00100 | DIST19 | PREFERRED BELOW | .00060 | DIST12 | PREFERRED ABOVE | .00100 |
| -.00050 TC | -.00010 | DIST17 | PREFERRED BELOW | -.00050 | DIST14 | PREFERRED ABOVE | -.00010 |
| -.00050 TC | -.00010 | DIST20 | PREFERRED BELOW | -.00050 | DIST14 | PREFERRED ABOVE | -.00010 |
| .00100 TC | .00500 | DIST16 | PREFERRED BELOW | .00100 | DIST15 | PREFERRED ABOVE | .00500 |
| .00100 TC | .00500 | DIST17 | PREFERRED BELOW | .00100 | DIST15 | PREFERRED ABOVE | .00500 |
| .00100 TC | .00500 | DIST18 | PREFERRED BELOW | .00100 | DIST15 | PREFERRED ABOVE | .00500 |
| .00010 TC | .00030 | DIST12 | PREFERRED BELOW | .00010 | DIST16 | PREFERRED ABOVE | .00030 |
| .00060 TC | .00100 | DIST17 | PREFERRED BELOW | .00060 | DIST18 | PREFERRED ABOVE | .00100 |
| -.00050 TC | -.00010 | DIST17 | PREFERRED BELOW | -.00050 | DIST20 | PREFERRED ABOVE | -.00010 |



^aThe branch to the left below a boundary interval is associated with the distribution unanimously preferred by decision makers less risk averse than the lower bound of the interval.

b. The branch to the right below a boundary interval is associated with the distribution unanimously preferred by decision makers more risk averse than the upper bound of the interval.

c. The bracketed intervals indicate the interval of absolute risk aversion consistent with the responses which lead to that branch on the hierarchy after the final choice is made.

Figure B.8 A Three-Stage Hierarchy of Questions

on the intervals at the center of the two segments of the measurement scale created by the first question, and the four questions at the third level focus on the boundary intervals at the center of the four segments created by the second set of questions.

Once the hierarchy of questions has been specified, the number of system output levels at which direct measurements of absolute risk aversion are to be made should be determined. Experience to date has shown that direct measurements in the neighborhood of three to four system output levels provide an adequate basis for the construction of an absolute risk aversion function over even a broad range of system output values. If, for example, annual income is the system output variable for which preference information is to be elicited and the relevant income range is from 0 to \$20,000, direct measurements of absolute risk aversion could be made in the neighborhood of \$3,000; \$10,000 and \$17,000.

In order to specify the choices used to elicit information on preferences in the neighborhood of a given system output level, the sample distributions generated by NORGEN must be shifted to that level by adding a constant to each element. The boundary intervals between distributions do not change when the expected value of their respective means is shifted away from zero. This is true because the reference levels on the measurement scale, λ , represent constant levels of absolute risk aversion, such as would be associated with a utility function of the form

$$u(y) = -e^{-\lambda y} \qquad \text{B.3}$$

When the mean of any distribution is shifted by adding a constant value to each of its elements, the associated expected utility is altered only

by a positive multiplicative factor for decision makers with this form of utility function.¹ If two distributions are shifted by the same amount, then, their relative ranking by decision makers more or less risk averse than any specified value remains unchanged.

A set of sample questions designed to elicit information on preferences for income in the neighborhood of \$10,000 is given in Figure B.9. They are specified in the manner described in Chapter IV. It should be noted that the choices actually presented to a decision maker are dependent upon his responses to prior questions in the hierarchy.

B.6 Administration and Interpretation of the Questionnaire

Before the questionnaire is administered, the decision maker should have a clear understanding of its objective, which is to obtain an accurate representation of his preferences. The system output for which preferences are to be measured should already have been clearly defined and should be recognized by the decision maker to be the primary indicator of system performance he will consider when making a choice in the situation being analyzed.

Administration of the questionnaire is straightforward. The decision maker is presented with several series of choices such as those

¹Let the mean of the distribution of a random variable y shift from zero to y^* --i.e., let the random variable w equal $y + y^*$.

$$\begin{aligned} E[u(w)] &= \int_{-\infty}^{\infty} e^{-\lambda w} dw \\ &= \int_{-\infty}^{\infty} e^{-\lambda(y+y^*)} dy \\ &= e^{-\lambda y^*} \int_{-\infty}^{\infty} e^{-\lambda y} dy \\ &= \alpha E[u(y)] \end{aligned}$$

where α is a positive constant.

1. Compare DIST 5 and DIST 20 and circle the one you prefer.

| | |
|--------|---------|
| DIST 5 | DIST 20 |
| | |

If you prefer DIST 20, go to question 3. Otherwise go to question 2.

2. Compare DIST 4 and DIST 7 and circle the one you prefer.

| | |
|--------|--------|
| DIST 4 | DIST 7 |
| | |

If you prefer DIST 7, go to question 5. Otherwise go to question 4.

3. Compare DIST 4 and DIST 8 and circle the one you prefer.

| | |
|--------|--------|
| DIST 4 | DIST 8 |
| | |

If you prefer DIST 8, go to question 7. Otherwise go to question 6.

4. Compare DIST 3 and DIST 20 and circle the one you prefer.

| | |
|--------|---------|
| DIST 3 | DIST 20 |
| | |

5. Compare DIST 3 and DIST 17 and circle the one you prefer.

| | |
|--------|---------|
| DIST 3 | DIST 17 |
| | |

6. Compare DIST 4 and DIST 20 and circle the one you prefer.

| | |
|--------|---------|
| DIST 4 | DIST 20 |
| | |

7. Compare DIST 4 and DIST 17 and circle the one you prefer.

| | |
|--------|---------|
| DIST 4 | DIST 17 |
| | |

Sample Distributions

| | | | |
|---------------|---------------|---------------|---------------|
| <u>DIST 3</u> | <u>DIST 4</u> | <u>DIST 5</u> | <u>DIST 7</u> |
| 9350 | 9550 | 9400 | 9000 |
| 9450 | 9700 | 9850 | 9550 |
| 9550 | 9800 | 9900 | 9850 |
| 9700 | 10050 | 9900 | 10400 |
| 10150 | 10100 | 10050 | 10450 |
| 10300 | 10200 | 10150 | 11100 |

| | | |
|---------------|----------------|----------------|
| <u>DIST 8</u> | <u>DIST 17</u> | <u>DIST 20</u> |
| 9050 | 9000 | 9250 |
| 9950 | 9050 | 9300 |
| 10000 | 9150 | 9600 |
| 10050 | 10000 | 10300 |
| 10150 | 10700 | 10400 |
| 10200 | 11100 | 10600 |

Figure B.9 Sample Questionnaire

specified in Figure B.8. Each series measures preferences in the neighborhood of a particular system output level. Completion of a questionnaire comprised of four three-question series takes approximately twenty minutes. Experience to date has shown that decision makers find this preference elicitation procedure more interesting and more informative than the interview process required to elicit a single-valued utility function.

Interpretation of the results is also quite straightforward. Using the series of questions in Figure B.9 as an example, consider the case in which the decision maker prefers DIST 5 in question (1), DIST 7 in question (4) and DIST 17 in question (5). Referring to Figure B.8, preference for DIST 5 over DIST 20 indicates that the decision maker is not less risk averse than .0001; i.e., that

$$r(y) > .0001 \quad \text{B.4}$$

Similarly, preference for DIST 7 over DIST 4 indicates that his level of absolute risk aversion is such that

$$r(y) < .0010 \quad \text{B.5}$$

Finally, from his preference for DIST 17 over DIST 3 it can be inferred that

$$r(y) < .0006 \quad \text{B.6}$$

As noted in the lower line of Figure B.8, then, these three responses indicate that the decision maker's level of absolute risk aversion lies on the interval $[\text{.0001}, \text{.0006}]$ in the neighborhood of $y = \$10,000$.

B.7 The Use of Interval Measurements of Preferences to Order Choices

The interval approach to the measurement of decision maker preferences was devised for use with the evaluative criterion of stochastic dominance with respect to a function. It determines upper and lower bounds on a decision maker's absolute risk aversion function, the basic information on preferences required for the application of their criterion. In order to actually implement stochastic dominance with respect to a function in the ordering of choices, however, utility functions having absolute risk aversion functions which correspond to these upper and lower bounds must be constructed. The link between absolute risk aversion functions and utility functions is straightforward, but analytical relationships between the two can be found only in certain special cases.¹ Because no particular functional form is specified for the upper and lower absolute risk aversion functions constructed under the interval approach, the determination of the associated utility functions by analytical means is a difficult if not impossible task.

Program UFUNC, which is listed in Figure B.10, resolves this problem. It employs numerical integration techniques to generate values for the utility functions associated with the upper and lower bound absolute risk aversion functions at regular intervals over the relevant range of system output levels. In effect, these values define the two

¹As Pratt (1964) notes,

$$u = \int e^{-\int r}$$

where u is a utility function and r is an absolute risk aversion function. The two constants of integration are arbitrary, corresponding to the arbitrary scale and origin of the utility function.

```

PROGRAM UFUNC(INPUT,OUTPUT,TAPE5=INPUT,TAPE4)
DIMENSION ARG(10),VALL(10),VALU(10)
DIMENSION U(2,400),W(400)
C READ MINIMUM AND MAXIMUM SYSTEM OUTPUT VALUES FOR WHICH
C UTILITY VALUES ARE TO BE CALCULATED, YMIN AND YMAX, THE SYSTEM
C OUTPUT INCREMENT DY, AND THE PRINT INTERVAL,YINT.
  READ(5,20) YMIN,YMAX,DY,YINT
98  FORMAT(4F10.2)
  NY=YINT/DY+.00001
  NK1=YMAX/YINT+.00001
  NK2=-YMIN/YINT+.00001
  IN=NK2+1
C READ THE NUMBER OF DATA POINTS FOR WHICH DIRECT INTERVAL
C MEASUREMENT INFORMATION IS TO BE GIVEN, KINT.
  READ(5,30) KINT
99  FORMAT(I5)
C READ PREFERENCE DATA.
  DO 2 I=1,KINT
    2  READ(5,100) ARG(I),VALL(I),VALU(I)
100  FORMAT(3F10.2)
C CONSTRUCT UPPER AND LOWER BOUND UTILITY FUNCTIONS.
  DO 50 MM=1,2
    C Y1 IS THE MARGINAL UTILITY AND Y2 IS THE LEVEL OF UTILITY.
    C THEY ARE SET AT 1.0 AND 0.0 RESPECTIVELY WHEN Y=0.0.
    Y1=1.0
    Y2=0.0
    Y=0.0
    IF(YMIN.EQ.0.0) GO TO 20
C THESE STATEMENTS DETERMINE UTILITY VALUES FOR NEGATIVE SYSTEM
C OUTPUT LEVELS.
    DO 30 J=1,NK2
      I=IND-J
      DO 5 K=1,NY
        Y=Y-DY
        IF(MM.EQ.1) RA=TABLI(VALL,ARG,Y,KINT)
        IF(MM.EQ.2) RA=TABLI(VALU,ARG,Y,KINT)
        Y2=Y2-DY*Y1
        Y1=Y1-DY*RA*Y1
5      CONTINUE
        U(MM,I)=Y2
        W(I)=Y
10     CONTINUE
        Y1=1.0
        Y2=0.0
        Y=0.0
20     I=NK2+1
        U(MM,I)=0.0
        W(I)=0.0
C THESE STATEMENTS DETERMINE UTILITY VALUES FOR POSITIVE SYSTEM
C OUTPUT LEVELS.
    DO 30 J=1,NK1
      I=I+1
      DO 25 K=1,NY
        Y=Y+DY
        IF(MM.EQ.1) RA=TABLI(VALL,ARG,Y,KINT)
        IF(MM.EQ.2) RA=TABLI(VALU,ARG,Y,KINT)
        Y2=Y2+DY*Y1
        Y1=Y1+DY*RA*Y1
25     CONTINUE
        U(MM,I)=Y2
        W(I)=Y
30     CONTINUE
50     CONTINUE
    KDIM=NK1+NK2+1
    DO 40 NN=1,KDIM
      PRINT 102 U(NN),U(1,NN),U(2,NN)
      WRITE(4,102) W(NN),U(1,NN),U(2,NN)
40     CONTINUE
102  FORMAT(IH ,F10.2,5X,2E20.13)
      END

```

Figure B.10 A Listing of Program UFUNC

utility functions, since values not calculated directly can be determined by linear interpolation, as in Figure B.11. Such an approximate representation of a function is called a table look-up function (Llewellyn, 1965).¹

Several inputs must be specified by the user of UFUNC. First, the range of system output levels over which the utility function is to be constructed is defined by specifying the minimum and maximum values, YMIN and YMAX. Values of the utility function over this range are calculated by solving the following two differential equations recursively with Euler integration:²

$$\frac{d}{dy} \begin{bmatrix} u(y) \\ u'(y) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -r(y) & 0 \end{bmatrix} \begin{bmatrix} u(y) \\ u'(y) \end{bmatrix} \quad \text{B.7}$$

The solution technique requires that initial values of $u(y)$ and $u'(y)$ be specified at some level of y . Within the program, $u(0)$ and $u'(0)$ are automatically set at 0 and 1.0 respectively, so this condition is met. It is also necessary to specify a value of DY , the output increment. The smaller the value of DY , the more accurate the numerical approximation of the utility function value will be. In cases where system output has been specified in dollars, values ranging up to 5.0 have proved to be adequate.³ If the range of systems outputs is large, more values of the

¹See Appendix A for a brief discussion of table look-up functions.

²Manetsch and Park (1977b) provide an excellent discussion of numerical integration in general and Euler integration in particular.

³Stability conditions under Euler integration require in this case that a value of DY be selected so that the absolute value of the following expression be less than 1.0 where R^* is the minimum value of the decision maker's absolute risk aversion function:

$$\frac{(2 - (DY) (R^*))}{2} \frac{(DY)^2 (R^*)^2 - 8(DY) (R)}{2}$$

If the minimum value of $r(y)$ for a decision maker is set at $-.01$, an extremely low level, the stability conditions are met if $DY = 5.0$.

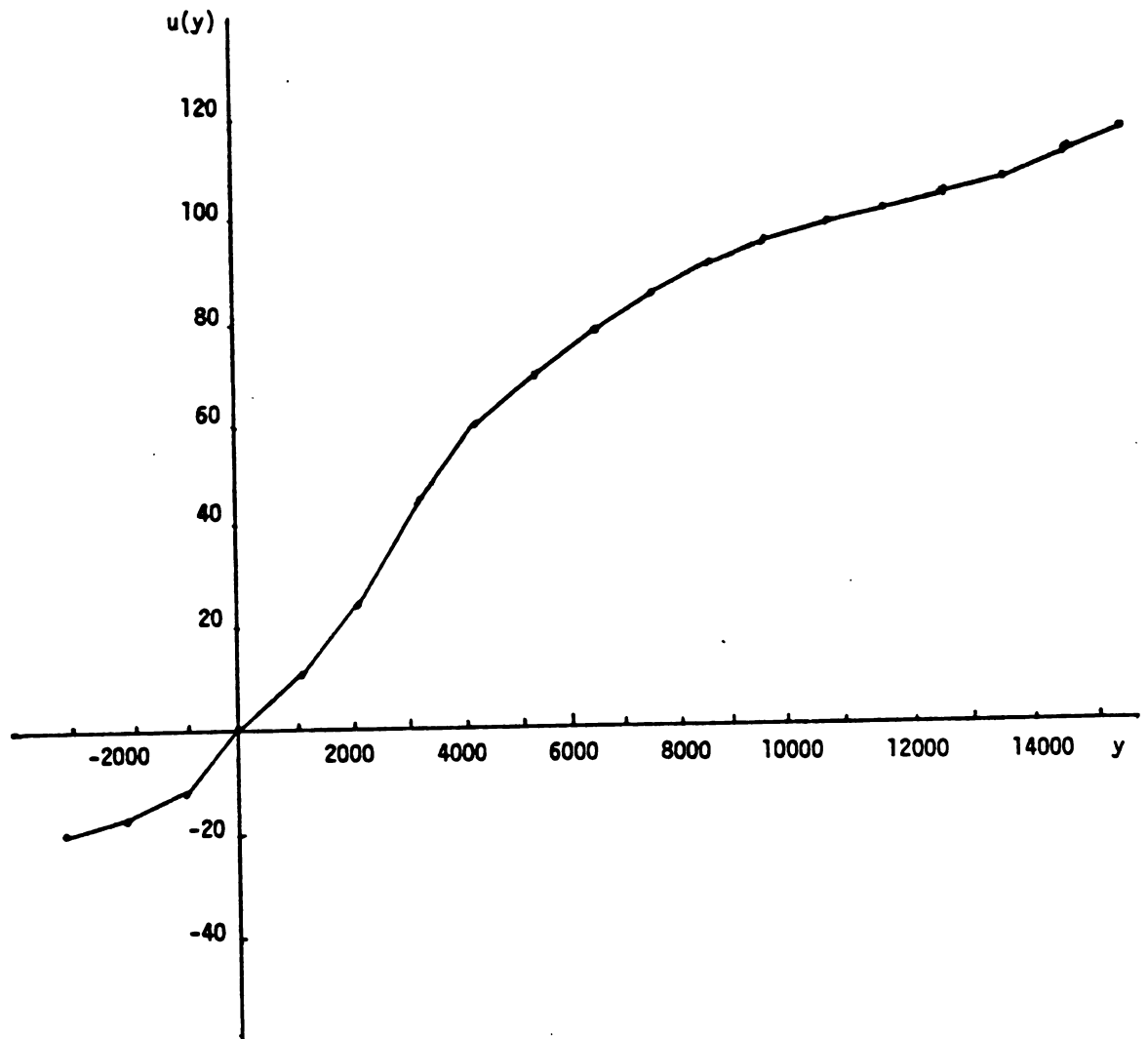


Figure B.11 Representation of a Utility Function by Interpolation Between Known Points

utility function will be calculated than are needed for the table look-up representation. A final parameter to be specified, then, is YINT. It defines the size of the interval between system output levels for which values of the utility function are to be specified in the table look-up function. In application to date, YINT has been set equal to 50 or 100. The minimum value for YINT is determined by the following expression

$$YINT > \frac{YMAX-YMIN}{399} \quad B.7$$

The user of UFUNC must also supply information on the decision maker's absolute risk aversion function. The interval approach to the measurement of preferences determines upper and lower bounds on a decision maker's absolute risk aversion function in the neighborhood of several system output levels. In the example shown in Figure B.12, direct interval measurements were made in the neighborhood of $y = 3000$, $y = 10,000$, and $y = 20,000$. The upper and lower bound functions are considered to be constant over the range of y values for which each measurement applies.¹ Values for the two absolute risk aversion functions at system output levels other than those where direct measurements have been made are determined by linear interpolation between known absolute risk aversion values or by linear

¹This is a result of the assumption that the decision makers absolute risk aversion function is constant in the neighborhood of any particular system output level (see Section 4.5 of Chapter IV). The range of system output levels over which a given measurement holds is dependent upon the dispersion of the sample distributions used to elicit the preference information. In the example in Figure B.12, distributions generated by NORGEN with STD set equal to 500 were used. As expected, nearly all points in the sample distribution fall within two standard deviations of the specified mean, y^* . Therefore, the interval measurements are said to be valid for system output levels in the range $y^* \pm 1000$.

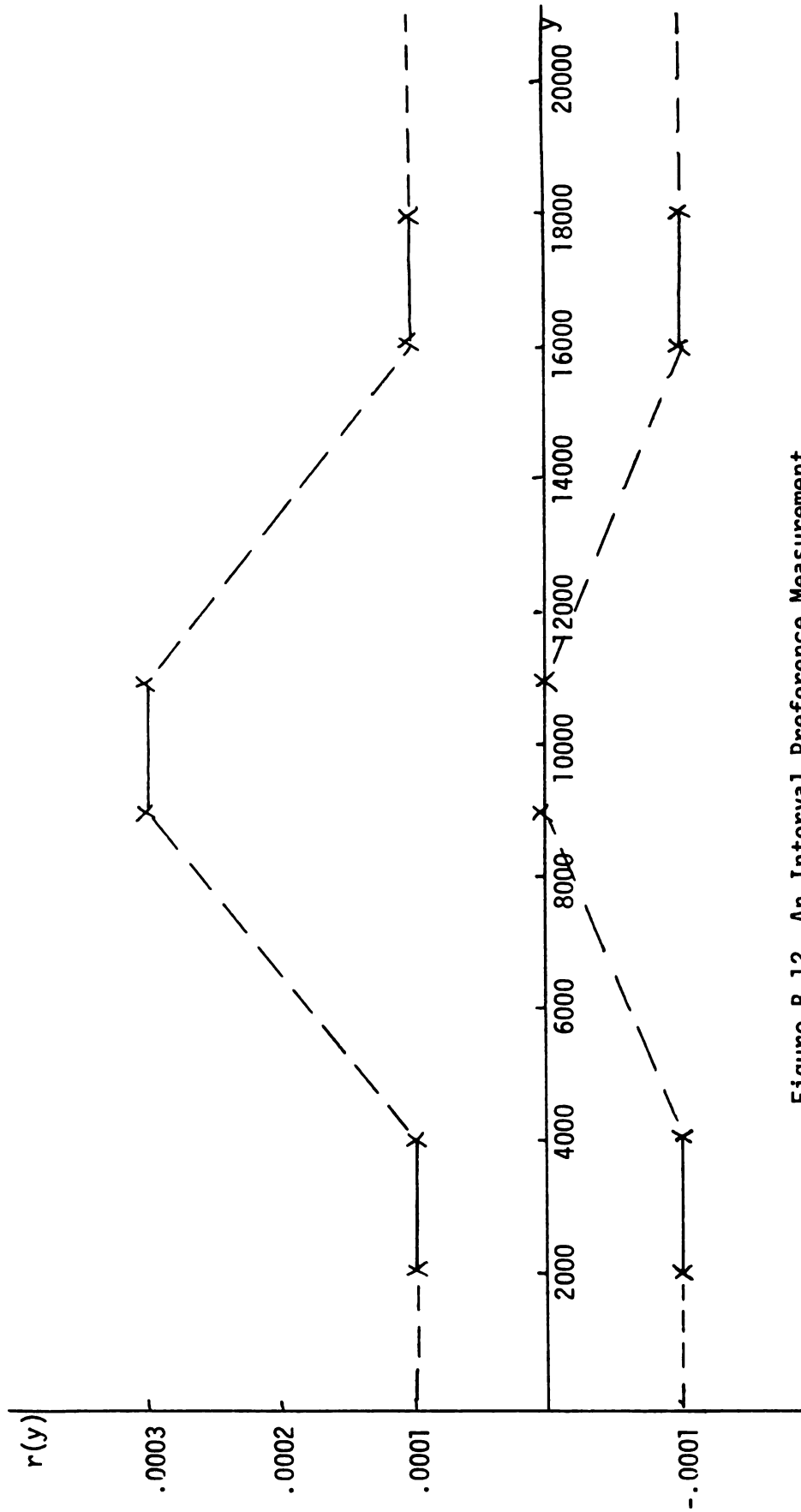


Figure B.12 An Interval Preference Measurement

extrapolation for system output levels outside the range over which direct measurements are made. The points marked with an X in Figure B.12 convey all the information on the upper and lower bound absolute risk aversion functions that is required by UFUNC. They occur at six system output levels. Therefore, parameter KINT is program UFUNC, which indicates the number of points for which specific information is required, is set equal to 6. A series of six data cards is then read by the program, each card setting values for three variables: ARG, which is the level of y ; VALL, which is equal to the lower value of $r(y)$; and VALU, which is equal to the upper value of $r(y)$. The values for the example in Figure B.12 are given in Table B.1. The utility functions generated by UFUNC for their example are graphed in Figure B.13.

Once values of the utility functions associated with the upper and lower bound absolute risk aversion functions have been calculated, they serve as inputs to program NSTDO, which orders distributions of system outputs according to the criterion of stochastic dominance with respect to a function. Program NSTDO is listed in Figure B.14.¹ The logical foundation of this procedure is explained in Section 4.4 of Chapter IV and, more extensively, in Meyer (1977a).

Several parameter values must be specified by the user of NSTDO. ND and NE again define the number of sample observations defining each system output distribution and the number of distributions to be considered. Their maximum values are 40 and 50 respectively. Data defining

¹This is a slightly modified version of the program written by Meyer for the application of stochastic dominance with respect to a function described in "Further Applications of Stochastic Dominance to Mutual Fund Performance."

Table B.1 An Example of Preference Data Input
for Program UFUNC

| ARG | VALL | VALU |
|-------|--------|-------|
| 2000 | -.0001 | .0001 |
| 4000 | -.0001 | .0001 |
| 9000 | 0 | .0003 |
| 11000 | 0 | .0003 |
| 16000 | -.0001 | .0001 |
| 18000 | -.0001 | .0001 |

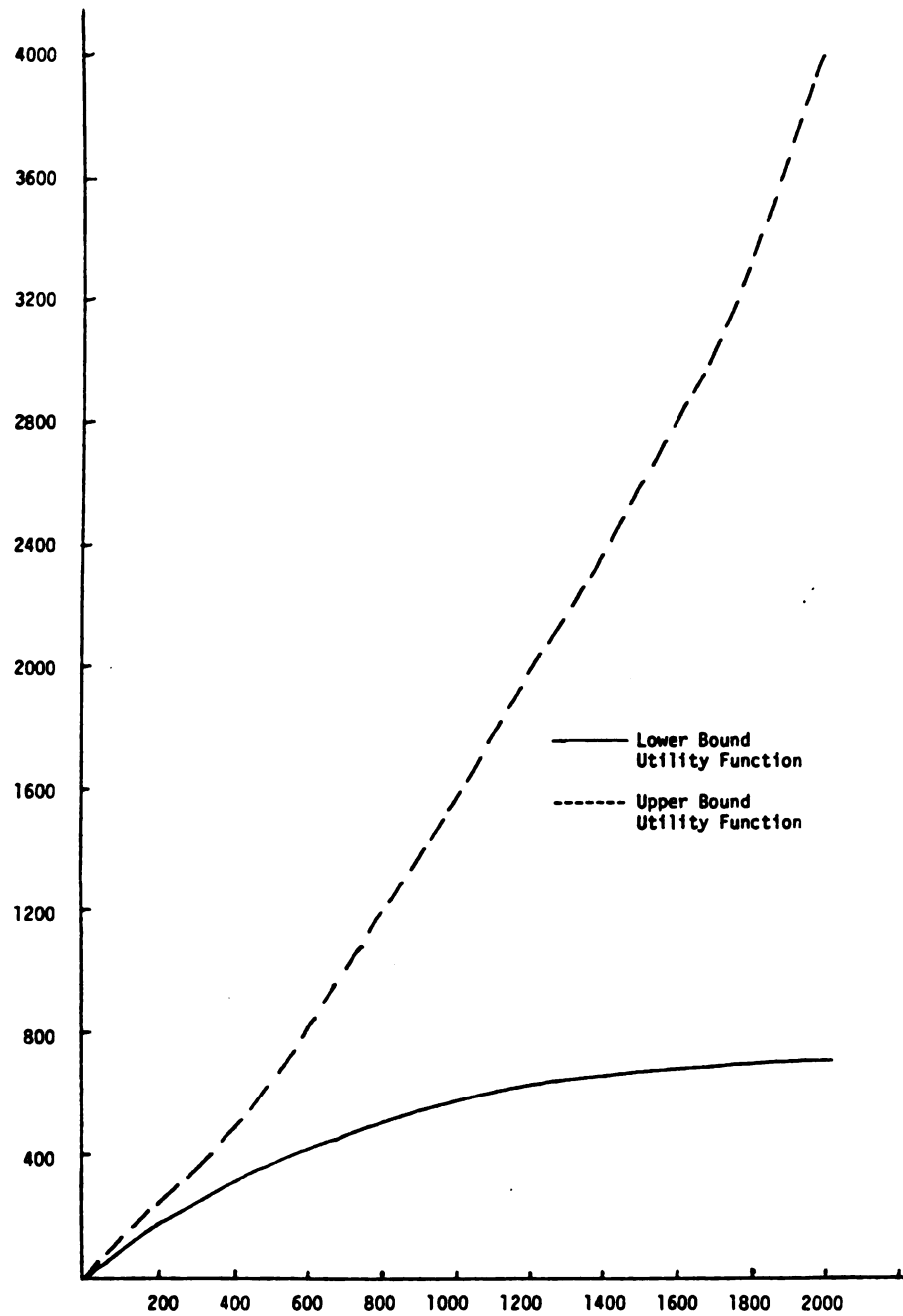


Figure B.13 Utility Functions Associated with Upper and Lower Bound Absolute Risk Aversion Functions

```

SUBROUTINE ROUNDER
COMMON ND,NE,ND2,C(101),CP(101),T(50),R(40,50),RA(10,4)
ND1=ND-1
DO 400 I=1,ND1
  AM=T(I)
  N=I
  II=I+1
  DO 401 J=II,ND
    IF(AM.LE.T(J)) GO TO 401
    N=J
  AM=T(J)
401 CONTINUE
  T(N)=T(I)
  T(I)=AM
400 CONTINUE
RETURN
END

SUBROUTINE COMCAL(M1,M2)
COMMON ND,NE,ND2,C(101),CP(101),T(50),R(40,50),RA(10,4)
K1=1
K2=1
C(1)=0.0
CP(1)=0.0
BND=ND
ND2=ND*2+1
DO 500 I=1,ND2
  IF(K1.GT.ND.AND.K2.GT.ND) GO TO 550
  IF(K1.GT.ND) GO TO 553
  IF(K2.GT.ND) GO TO 551
  IF(R(M1,K1)-R(M2,K2)) 551,552,553
551 C(I)=C(I-1)+(1.0/BND)
  CP(I)=R(M1,K1)
  K1=K1+1
  GO TO 500
552 C(I)=C(I-1)
  CP(I)=R(M1,K1)
  K1=K1+1
  K2=K2+1
  GO TO 500
553 C(I)=C(I-1)-(1.0/BND)
  CP(I)=R(M2,K2)
  K2=K2+1
  GO TO 500
550 C(I)=0.0
  CP(I)=0.0
500 CONTINUE
RETURN
END

SUBROUTINE SD(MU,II)
COMMON ND,NE,ND2,C(101),CP(101),T(50),R(40,50),RA(10,4)
EXTERNAL U1,U2
REAL MU
MU=0.0
NL=2*ND
ND2=2*ND
DO 600 I=1,ND2
  IF(NL.LE.1) GO TO 650
  IF(C(NL)) 651,652,653
652 NL=NL-1
  MU=0.0
  GO TO 600
653 MU=MU+(U2(CP(NL+1),II)-U2(CP(NL),II))*C(NL)
  NL=NL-1
  DO 601 J=1,ND2
    IF(NL.LE.1) GO TO 650
    TU=(U2(CP(NL+1),II)-U2(CP(NL),II))*C(NL)
    IF(MU+TU) 654,652,655
655 MU=MU+TU
    NL=NL-1
601 CONTINUE
654 TT=MU*(C(NL)+U2(CP(NL+1),II)
  Q=U2(TT,II)
  MU=0.0
  CP(NL+1)=Q
  GO TO 600
651 MU=MU+(U1(CP(NL+1),II)-U1(CP(NL),II))*C(NL)
  NL=NL-1
  DO 602 K=1,ND2
    IF(NL.LE.1) GO TO 650
    TU=(U1(CP(NL+1),II)-U1(CP(NL),II))*C(NL)
    IF(MU+TU) 656,652,657
656 MU=MU+TU
    NL=NL-1
602 CONTINUE
657 TS=MU*(C(NL)+U1(CP(NL+1),II)
  Q=U1(TS,II)
  MU=0.0
  CP(NL+1)=Q
600 CONTINUE
650 RETURN
END

```

Figure B.14 A Listing of Program NSTD0^a^aThis is a modified version of a program written by Jack Meyer.

Figure B.14 (cont'd.)

```

PROGRAM NSTDD(INPUT, OUTPUT, TAPE5=INPUT, TAPE6=OUTPUT, TAPE4)
COMMON ND, NE, NG, C(101), CP(101), T(50), R(40, 50), RA(10, 4)
COPEN(1, 2) DCKRZ ARG(400), VALL(400), VALU(400), KDIM, SMALL, DIFF
DITERATION, ICEBE(40, 3), NSD(50), CE1(40, 50), CE2(40, 50), NAME(40)
READ(5, 900) NE, NE
FORMAT(2F10)
950  WRITE(6, 300)
    NG=1
    WRITE(6, 300)
300  FORMAT(1H1, 7X, 5HNAME, 21X, 11HDATA POINTS, 23X, 16HMEAN STD)
    DO 100 I=1, NE
    READ(5, 1000) NAME(I)
4000  FORMAT(1A10)
    DO 3000 MME=1, ND
    READ(5, 1000) R(I, MME)
5000  CONTINUE
5001  FORMAT(1F10, 2)
    VALL(I)=0
    AMEAN=0
    BND=ND
    DO 101 J=1, ND
    AMEAN=AMEAN+(1.0/BND)*R(I, J)
    VALL=VALL+(1.0/BND)*R(I, J)*R(I, J)
    T(I)=T(I, J)
101  CONTINUE
    VARSQRT=AMEAN*AMEAN
    STD=SQRT(VARSQRT)
    WRITE(6, 400) NAME(I), (R(I, J), J=1, 10)
    WRITE(6, 401) (R(I, J), J=11, 20), AMEAN, STD
400  FORMAT(1H, A10, 10(1X, F8.2))
401  FORMAT(1H, 10X, 10(1X, F8.2), 2(4X, F8.2))
    CALL RORDER
    DO 102 J=1, ND
    R(I, J)=T(J)
102  CONTINUE
100  CONTINUE
    WRITE(6, 300)
305  FORMAT(1H1)
    DO 103 I=1, NG
    READ(5, 900) SMALL, DIFF, KDIM
900  FORMAT(2F10, 2, 15)
    DO 800 MME=1, KDIM
    READ(5, 901) ARG(MME), VALL(MME), VALU(MME)
901  FORMAT(1H, F10, 2, 5X, 2E20, 13)
800  CONTINUE
    DO 104 J=1, NE
    B1=0
    B2=0
    DO 105 K=1, ND
    B1=B1+(1.0/BND)*R(J, K), I
    B2=B2+(1.0/BND)*R(J, K), I
105  CONTINUE
    CE1(J, 1)=B1(1)
    CE2(J, 1)=B2(1)
104  CONTINUE
103  CONTINUE
    WRITE(6, 303)
303  FORMAT(1H1, 10X, 4HNAME, 6X, 6HVERSUS, 6X, 4HNAME, 31X, 6HGROUPS)
    DO 115 I=1, NE
    DO 106 J=1, NE
    IF(J.EQ.I) GO TO 106
    DO 107 K=1, NG
    IF(CE1(I, K).EQ.CE1(J, K).AND.CE2(I, K).EQ.CE2(J, K)) GO TO 60
    IF(CE1(I, K).LE.CE1(J, K).AND.CE2(I, K).LE.CE2(J, K)) GO TO 61
    NSD(K)=0
    GO TO 107
60  CALL CUNVAL(J, I)
    CALL EDCEU(K)
    IF(CEU.GE.0.0) GO TO 62
    IF(CE1(I, K).EQ.CE1(J, K).AND.CE2(I, K).EQ.CE2(J, K)) GO TO 61
    NSD(K)=0
    GO TO 107
62  NSD(K)=1
    GO TO 107
61  CALL CUNVAL(I, J)
    CALL SKEU(K)
    IF(CEU.GE.0.0) GO TO 63
    NSD(K)=0
    GO TO 107
63  NSD(K)=-1
107  CONTINUE
    WRITE(6, 304) NAME(I), NAME(J), (NSD(K), K=1, NG)
304  FORMAT(1H, 6X, A10, 10X, A10, 2X, 10I6)
106  CONTINUE
114  WRITE(6, 306)
306  FORMAT(2)
115  CONTINUE
    STOP
    END)

```

Figure B.14 (cont'd.)

```

FUNCTION UI(X,JJ)
COMMON ND,NE,NC,C(101),CP(101),T(50),R(40,50),RA(10,4)
COMMON /BLOCKB/ ARG(400),VALL(400),VALU(400),KDIM,SMALL,DIFF
DUMMY=X
K=KDIM-1
UI=TABEX(VALL,SMALL,DIFF,K,DUMMY)
RETURN
END

```

```

FUNCTION UP(X,JJ)
COMMON ND,NE,NC,C(101),CP(101),T(50),R(40,50),RA(10,4)
COMMON /BLOCKB/ ARG(400),VALL(400),VALU(400),KDIM,SMALL,DIFF
DUMMY=X
K=KDIM-1
UP=TABEX(VALU,SMALL,DIFF,K,DUMMY)
RETURN
END

```

```

FUNCTION UI1(X,JJ)
COMMON ND,NE,NC,C(101),CP(101),T(50),R(40,50),RA(10,4)
COMMON /BLOCKB/ ARG(400),VALL(400),VALU(400),KDIM,SMALL,DIFF
K=KDIM
UI1=TABEX(ARG,VALL,X,K)
RETURN
END

```

```

FUNCTION UI2(X,JJ)
COMMON ND,NE,NC,C(101),CP(101),T(50),R(40,50),RA(10,4)
COMMON /BLOCKB/ ARG(400),VALL(400),VALU(400),KDIM,SMALL,DIFF
K=KDIM
UI2=TABEX(ARG,VALU,X,K)
RETURN
END

```

```

FUNCTION TABEX(VALL,SMALL,DIFF,K,DUMMY)
DIMENSION VAL(1)
DUM=DUMMY-SMALL
I=MINO(MAXI(1.0+DUM/DIFF,1.0),K)
TABEX=VAL(I+1)-VAL(I)*(DUM-FLOAT(I-1)*DIFF)/DIFF+VAL(I)
RETURN
END

```

```

FUNCTION TABEX(VAL,ARG,DUMMY,K)
DIMENSION VAL(1),ARG(1)
DO 1 J=2,K
IF(DUMMY-ARG(J)) GO TO 1
2 TABEX=(DUMMY-ARG(J-1))*(VAL(J)-VAL(J-1))/
1 (ARG(J)-ARG(J-1))+VAL(J-1)
RETURN
1 CONTINUE
J=K
GO TO 2
END

```

the NE distributions are read next by the program. This information is stored in arrays NAME and R. Finally, data on the decision maker's preferences must be read into the program. Values of SMALL, DIFF, and KDIM--the smallest system output value for which utility values are to be assigned, the difference between system output levels for which utility values are to be assigned, and the number of system outputs for which utility values are to be assigned--are read first. Then data on the utility functions associated with the decision maker's upper and lower bound absolute risk aversion functions are read into arrays ARG, VALL, and VALU, which are defined as above in program UFUNC.¹

A simple output from program NSTDO is shown in Figure B.15. This is an ordering of four distributions for the decision maker whose interval preference measurement is graphed in Figure B.12. The symbol 1 indicates that the first distribution named is preferred to the second; -1 indicates that the second distribution is preferred to the first; and 0 indicates that the two distributions cannot be ordered by the criterion of stochastic dominance with respect to a function for the class of decision makers whose absolute risk aversion functions lie within the specified bounds.

The combined power of interval measurements of decision maker preferences and the criterion of stochastic dominance with respect to a function is demonstrated by the results presented in Sections 4.7 and 4.8 of Chapter IV. Clearly these are two related analytical tools which can be of considerable value in the analysis of decision made under uncertainty.

¹The program reads these data from a permanent file which is the catalogued output of program UFUNC.

| NAME | VERSUS | NAME | |
|--------|--------|--------|----|
| DIST 1 | | DIST 2 | 0 |
| DIST 1 | | DIST 3 | 0 |
| DIST 1 | | DIST 4 | -1 |
| DIST 2 | | DIST 1 | 0 |
| DIST 2 | | DIST 3 | -1 |
| DIST 2 | | DIST 4 | -1 |
| DIST 3 | | DIST 1 | 0 |
| DIST 3 | | DIST 2 | 1 |
| DIST 3 | | DIST 4 | -1 |
| DIST 4 | | DIST 1 | 1 |
| DIST 4 | | DIST 2 | 1 |
| DIST 4 | | DIST 3 | 1 |

Figure B.15 A Sample Output from Program NSTD0

In this appendix the relatively straightforward procedures used to implement these techniques in a practical setting have been described.

APPENDIX C

IMPLEMENTATION OF THE GREMP MODEL

C.1 Introduction

In this appendix the computer program used to implement the GREMP model is described, and some of the special features of the model are discussed. The objective is to acquaint the potential user with the more technical aspects of this procedure and to suggest ways in which the model can be adapted for use in the analysis of particular decision problems.

The basic structure of the computer program which implements the GREMP model is shown in the flow chart in Figure C.1. The program begins with an initialization phase, during which parameter values are specified and required data are read in. The program then goes through a specified number of iterations during which strategies are generated at random, the outcomes of each strategy are simulated for a number of states of nature, and the efficient set is updated. Once the desired number of alternative strategies has been generated and evaluated, information on the elements in the efficient set is printed, and the program terminates.

As was noted in Chapter V, this procedure is not designed to identify a truly optimal choice. Rather, it simply generates a large number of strategies, and, on the basis of evaluative information supplied by the user, it identifies an efficient set of choices from those considered. The particular value of the GREMP model is that it can be used to analyze problems for which an optimal solution cannot be determined analytically.

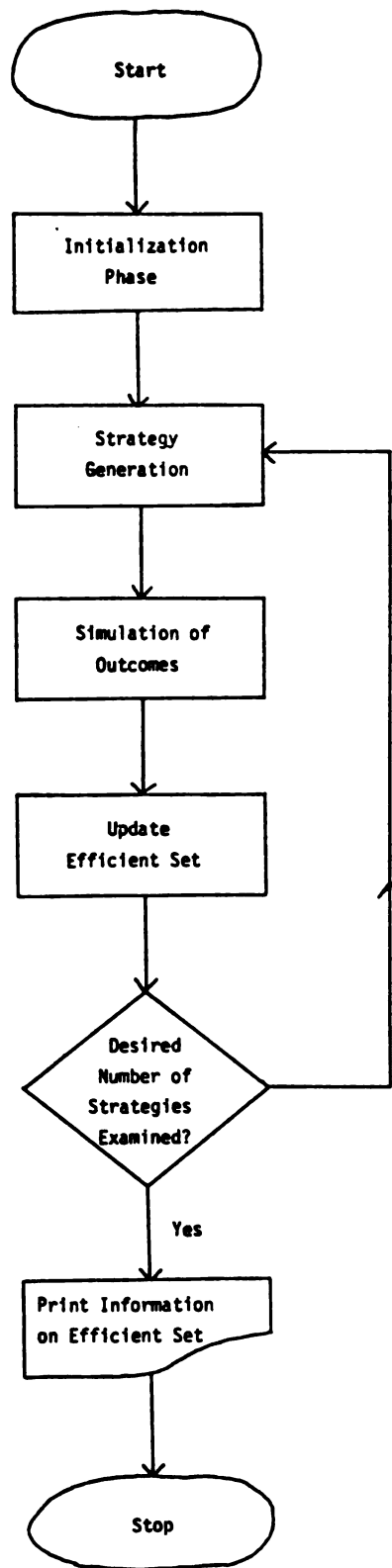


Figure C.1 General Flow Chart of Program GREMP

The discussion in subsequent sections is organized in a manner similar to the computer program itself. Data requirements and suggested program parameter values are first examined. Next, the procedure by which strategies are generated is described. The simulation of the outcomes associated with each strategy generated and the evaluation of strategies by the criterion of stochastic dominance with respect to a function are then briefly discussed, with references being made to the more extensive descriptions of these procedures given in Chapters III and IV and in Appendices A and B. A complete listing of the program is included at the end of the Appendix.

C.2 The Initialization Phase

During the initialization phase of the program, run parameter values are established, some or all of the constraints on control variable levels are specified, and data defining alternative states of nature and the decision maker's preferences are read in.

The run parameters define certain general characteristics of any particular application of the model. They include: ND, ITNS, NV, NC, NVC, MAXNO and NCONS. As in other programs developed in this study, ND is the number of sample observations defining the distribution of outcomes associated with each strategy being considered. As such, it is also the number of states of nature to be defined. In the applications of the GREMP model discussed in Chapters V and VI, a value of 20 was specified for ND. In many practical instances a larger value of ND would be desirable. The maximum value in this version of the program is 20, but this can be augmented by simply changing the appropriate

array dimensions.¹ ITNS is the number of iterations the model will perform--i.e. the number of sample strategies which will be generated and evaluated. The value of ITNS depends entirely on the characteristics of the problem being analyzed. ITNS was set at 500 and 1000 in the two applications of the GREMP model discussed in this study. It may be desirable to specify larger values of ITNS when problems with a large number of choice variables are to be analyzed or when the identification of a more nearly optimal strategy is desired. If the simulation model used to generate sample observations from the distribution of outcomes associated with each strategy is quite complex, however, the cost of each iteration may be so high that a much lower value of ITNS must be specified.

NV is the number of control variables used to define a management strategy in the problem being considered. The set of control variables can be divided into as many as ten categories, with NC being the number of categories. In the applications discussed in Chapters V and VI, for example, it was convenient to divide the control variables into three categories: resource acquiring activity levels, resource using activity levels, and control rule parameters.² Once the number of categories has been specified, the number of control variables in each category must be

¹Each stochastic system input variable must be dimensional to ND; e.g. in the listing at the end of this chapter, the second argument in each array in common blocks 2 and 3 is set to ND. In addition, T and R are dimensioned to T(ND) and R(21,ND), and C and CP are dimensioned to C(2*ND+1) and CP(2*ND+1).

²See Eisgruber and Lee (1971) for an interesting discussion of why choice variables need to be classified in such a manner when strategies are to be constructed in a sequential manner.

read into the appropriate element of the array NVC--i.e. NVC(1) is set equal to the number of control variables in category one; NVC(2) is set equal to the number in category two; etc. Since the classification of variables is intended to be mutually exclusive and exhaustive, the following equation must hold:

$$NV = \sum_{i=1}^{NC} NVC_i \quad C.1$$

In some instances it may be infeasible for all NV of the control variables to be set at non-zero levels in the specification of a strategy. When this is the case, it may be desirable to impose a limit on the number of control variables considered in defining each strategy. MAXNO is used to impose such a restriction; its value must be less than or equal to NV.¹

Finally, NCONS is the number of linear constraints to be imposed on the control variables which define a management strategy. In general, all these constraints must be of a "less than or equal to" form.² Up to 25 linear constraints can be imposed in the current version of the model. This number can easily be expanded, however.

Once the run parameter values have been specified, the program reads information on two sets of constraints. Members of the first set limit the range of allowable values for the NV control variables by establishing a minimum value, VMIN, a maximum value, VMAX, and the magnitude of the interval between values, VINT. If, for a particular

¹See Donaldson and Webster (1968) for a more complete discussion of how such restrictions can be used.

²See Dent and Thompson (1968) for a discussion of the difficulties caused by constraints of other forms and for an explanation of how these difficulties can be overcome.

control variable, $V_{MIN} = 0$, $V_{MAX} = 100$, and $V_{INT} = 25$, then possible values for that variable are 0, 25, 50, 75, and 100. If V_{INT} is set equal to an integer, all values of the control variable will be integer. If, on the other hand, V_{INT} is set equal to a very small value, the set of allowable values approaches that of a continuous variable between V_{MIN} and V_{MAX} .

Linear constraints such as those used in the specification of a linear programming model comprise the second set of constraints. The program reads input-output coefficients and resource availability levels for each of the $NCONS$ constraints of this type. All input-output coefficients are first initialized to equal zero. Non-zero values are then read into the two dimensional array A by specifying the constraint number, I , the control variable number, J , and the desired value of $A(I,J)$.¹ The variable $LAST$ is simply a flag which, when set to a non-zero value, indicates that the last non-zero input-output coefficient has been read. Next, the $NCONS$ resource availability levels are read into the array F , and the specification of linear constraints is completed.

The initialization phase continues with the program reading data which define levels for each stochastic exogenous system input variable in each of the ND states of nature used in the determination of system output distributions.² The number of exogenous system input

¹Control variables are ordered in the following manner. The first $NVC(1)$ control variables are the elements in category one, the next $NVC(2)$ control variables are the elements in category two, etc.

²These data are generated externally to the program using techniques described in Appendix A.

variables depends on the characteristics of the system being considered in a particular decision analysis. In general, then, the user will supply his own READ statements here. The version of the program listed at the end of this appendix is that used in the analysis of the problem discussed in Chapter VI. Therefore, data on contract prices, cash prices, crop yields, and days available for fieldwork are read.¹

Finally, information on decision maker preferences required for the application of stochastic dominance with respect to a function is read by the program. More specifically, the data points generated by program UFUNC, which define the utility functions associated with the decision maker's upper and lower bound absolute risk aversion functions, are read.² First, however, values of YMIN, DY, and KDIM--parameters of the table look-up functions used to represent those utility functions--are read. YMIN is the minimum value of the system output variable for which a utility value is calculated, DY is the interval between system output levels for which utility values are calculated, and KDIM is the total number of data points. Once these values are established, KDIM values of ARG, VALL, and VALU--a system output level, a lower bound utility value, and an upper bound utility value--are read into the appropriate arrays.³

¹All these data are read from separate permanent files. This is often more convenient than using data cards.

²See Appendix B for a listing and description of this program.

³These values are read from a permanent file which is the catalogued output of UFUNC.

C.3 Strategy Generation

All user supplied inputs are read into the program during the initialization phase. At the beginning of the first iteration and of all subsequent iterations, all control variables are set to zero, all constraints are reset to their original values, and several variables used to monitor the strategy generation process are set to zero. The generation of a feasible strategy then begins.

The segment of the main program which constructs each strategy is listed in Figure C.2. The sequence of operations is such that all the elements of one control variable category are assigned values before those in the next category are considered. After updating values of ILO and IHI, the lowest and highest variable numbers of the elements in the variable category being considered, the program calls subroutine SELECT.¹ This subroutine is a discrete uniform process generator which randomly selects a variable, $V(J)$, from the set of variables in the category under consideration. If that variable has been considered in the construction of the current strategy, the value of $IND(J)$ will be non-zero and subroutine SELECT will be called again. If it has not been considered, the value of $NEX(I)$, the number of variables within the category already examined, is augmented by one and $IND(J)$, the indicator for the variable is set equal to $NEX(I)$. Subroutine LEVSET is then called. Like SELECT, it is a discrete uniform process generator.

¹When none of the control variables in a particular category imposes constraints on any of the others, it is possible to bypass calls of subroutines SELECT and CHECK. In the version of the program listed at the end of this appendix, this is done for variable category 3, the set of control rule parameters.


```

      ILO=1
      IHI=0
      DO 65 I=1,NC
      ILO=IHI+1
50    CALL SELECT(ILO,IHI,J)
      IF(IND(J) NE 0) GO TO 50
      NEX(I)=NEX(I)+1
      IND(J)=NEX(I)
      CALL LEVSET(J)
      CALL CHECK(J,NCUNS,NIN)
      IF(NIN GE MAXND) GO TO 70
      IF(NEX(I) LT IND(I)) GO TO 50
65    CONTINUE
70    CALL EXPAND

```

Figure C.2 Statements in GREMP Which Generate Feasible Strategies

Given values of VMIN, VMAX, and VINT for variable J, it sets a level for that control variable.¹

Once a value for a particular control variable has been set, the feasibility of that value must be checked. This is done by calling subroutine CHECK. If all the constraints imposed in the management strategy are linear and of a "less than or equal to" form, and if all values of the array A in column J are of the same sign, CHECK is as listed in Figure C.3. The feasibility of the specified level V(J) is checked against each of the NCONS constraint levels. If a constraint is violated, the value of V(J) is adjusted to a level which is feasible. Once all constraints have been met, each constraint level is updated to reflect the resource requirements associated with the level of V(J) and control is returned to the main program.

In many instances, there may be a need to impose non-linear constraints on some of the control variables. One common non-linear constraint takes the form:

$$V(J)*V(K) = 0, \quad \text{C.2}$$

which implies that V(J) and V(K) cannot both have non-zero values. Such a constraint can easily be incorporated into subroutine CHECK by adding statements such as the following:

```
IF(V(K).NE.0.0) V(J) = 0.0
IF(V(J).EQ.0.0) GO to 20
V(K) = 0.0
NEX(2) = NEX(2)+1
IND(k) = NEX(2)
C.3
```

¹As was noted in the preceding section, if VINT is set to a very small value, the variable V becomes, for all practical purposes, a continuous variable. So LEVSET can approximate a continuous uniform process generator.

```

SUBROUTINE CHECK(J,NCONS,NIN)
COMMON /BLOCK1/  V(25),VMIN(25),VMAX(25),VINT(25)
COMMON /BLOCK2/  A(25,25),F(25),IND(25)
IF(V(J).EQ.0.0) GO TO 20
DO 10 I=1,NCONS
IF(A(I,J).EQ.0.0) GO TO 10
IF(F(I).GE.A(I,J)*V(J)) GO TO 10
LEV=TV/VINT(J)
V(J)=FLOAT(LEV)*VINT(J)
IF(V(J).LT.VMIN(J)) V(J)=0.0
IF(V(J).GT.VMAX(J)) V(J)=VMAX(J)
10 CONTINUE
IF(V(J).EQ.0.0) GO TO 20
NIN=NIN+1
DO 15 I=1,NCONS
15 F(I)=F(I)-V(J)*A(I,J)
20 RETURN
END

```

Figure C.3 A Simple Version of Subroutine CHECK

If $V(K)$ already has a non-zero value, $V(J)$ is set equal to zero and control is returned to the main program. If $V(J)$ is not equal to zero, $V(K)$ is set equal to zero. The number of variables in its category--in this case variable category two--is augmented by one and the indicator for this variable, $V(K)$, is set equal to $NEX(2)$. These final two statements ensure that $V(K)$ will not be considered again later in the strategy generation process. Other types of non-linear constraints can be treated in a similar manner. Examples of constraints which closely parallel that given in Equation C.2 include:

$$V(J) * V(K) \geq F(I) \quad C.4$$

$$V(J) / V(K) \leq F(I) \quad C.5$$

and

$$\text{if } V(J) \geq F(I), \text{ then } V(K) = 0 \quad C.6$$

Subroutine CHECK can also be adapted to make the strategy generation process more efficient in some instances. For example, if the only constraint on two variables $V(J)$ and $V(K)$ is

$$V(J) + V(K) = F(I), \quad C.7$$

then a value of $V(K)$ can be established automatically once a value of $V(J)$ has been specified and vice versa. Still another useful alteration of subroutine CHECK in some situations involves the addition of statements which alter levels of V_{MIN} or V_{MAX} for all or some variables in a category to reflect the impact of the level selected for another variable. This was done in the applications presented in this study when maximum allowable acreages for corn and soybeans were adjusted upward to reflect additions in available acreage due to land rental.

The strategy generation process continues until all the variables in each category have been considered or until the maximum number of

variables have entered the control strategy at a non-zero level. At this point, however, it is possible that the strategy constructed does not use the available resources to the fullest extent. Therefore, it may be desirable to expand the strategy out until one or more additional constraints become binding. This is done by subroutine EXPAND.

Donaldson and Webster (1968) provide a good discussion both of the rationale for expanding strategies and of the means by which this can be done. It is difficult, however, to specify a general form for subroutine EXPAND. In most instances its form is problem-specific and so this subroutine must be supplied by the user. In the listing presented at the end of this appendix, EXPAND is used to increase acreage levels of corn or soybeans, depending on which was specified first, so that all available acreage is used.

C.4 The Simulation of Strategy Outcomes

When the strategy generation phase of each iteration has been completed, subroutine DISGEN is called. This user-supplied subroutine simulates system performance under the newly specified strategy for each of the ND states of nature. Subroutine DISGEN can be as simple or as complex as the problem under analysis requires. It can be as simple as the example listed in Figure C.4, in which the level of system output realized is a simple linear function of five control variables and of the stochastic price, yield, and cost levels associated with each. Alternatively, DISGEN can be as complex or still more complex than that in the version of the program listed at the end of this chapter.

```

SUBROUTINE DISGEN
COMMON /BLOCK1/ P(5,20),Y(5,20),C(5,20)
COMMON /BLOCK2/ V(5)
COMMON /BLOCK3/ ND,NP,R(21,20)
DO 5 J=1,ND
R(NP,J)=0.0
DO 5 I=1,5
5 R(NP,J)=R(NP,J)+(P(I,J)*Y(I,J)-C(I,J))*V(I)
RETURN
END

```

Figure C.4 A Simple Version of Subroutine DISGEN

A brief comment should be made about the variables passed to and from DISGEN. First, it should be noted that R, the array of system output levels passed back to the main program, has two dimensions. The first argument of R is always set equal to NP in subroutine DISGEN. This parameter identifies the system output distribution during the evaluative phase of each iteration. NP is passed to the subroutine and should be passed back unchanged. Values of all control variables--the array V--and all exogenous system input levels are also passed to and from the subroutine unchanged.

Finally, it should also be noted that there is no reason why the distribution for more than one system output variable cannot be determined with each call of DISGEN. The evaluative criterion used in the GREMP model, stochastic dominance with respect to a function, can, however, only evaluate alternative strategies by considering the distributions of a single system output variable.

C.5 Updating the Efficient Set

During the final phase of each iteration the efficient set is updated by comparing the system output distribution of the newly generated strategy to the distribution of each strategy in the efficient set. This is done by subroutine NSTDO, which is simply a modified version of the program written by Meyer to implement the criterion of stochastic dominance with respect to a function. This program and the required inputs to it were described in the final section of Appendix B. Those same inputs must be passed to subroutine NSTDO from the main program. They include values for ND, the number of sample observations defining each system output distribution and values in the array R, the

array of simple system output levels for each plan in the current efficient set and for the newly generated plan. Information in decision maker preferences is passed from the main program directly to function subprograms U1, U2, UI1, and UI2, which are called by subroutine NSTDO.

As described in Section 5.3.3 of Chapter V the evaluative process involves the pair-wise comparison of the distribution of system output levels associated with the newly generated strategy with that associated with each strategy in the efficient set. Such comparisons are made until the new strategy is dominated or until all possible comparisons have been made. If the new strategy is dominated, control is returned immediately to the main program. If it dominates strategies in the efficient set, those which are dominated are eliminated and the arrays describing the efficient set--R, AMEAN, STD, and VLEV--are rearranged. The program limits the size of the efficient set to twenty strategies. This is done because it reduces storage requirements and because efficient sets having more than twenty elements are difficult to work with. When the size of the efficient set reaches the maximum allowable level, newly generated strategies are still compared to each element in the efficient set, but they enter the efficient set only if they dominate one or more of its existing elements.

C.6 Output of the GREMP Model

Information on selected run parameter values, on the linear constraints specified, and in the range of allowable values for each control variable is printed before the first iteration is begun. This information can be of value for diagnostic purposes. When the final iteration has been completed, information on each element in the efficient set is

printed. First, the mean and standard deviation of the system output distribution associated with a particular strategy are printed, then the control variable levels defining the strategy are printed, and, finally, the set of sample observations defining the system output distribution is printed. It should be noted that performance indicators other than those upon which the evaluation of alternatives is based could also be printed if this were desired.

C.7 A Listing of the GREMP Model

```

PROGRAM GREMP(INPUT,CUTPUT,TAPFF=INPUT,TAPE4,TAPE6,TAPE7,TAPE8,
1TAP9)
COMMON /FLOCK1/ V(25),VMIN(25),VMAX(25),VINT(25)
COMMON /FLOCK2/ A(25,25),F(25),IN(25)
COMMON /FLOCK3/ P(5,25),V(25,25),E(5,11,20)
COMMON /FLOCK4/ CFC(12,20),EYC(12,20),CFS(12,20),CYS(12,20)
COMMON /FLOCK5/ NRC(5,5),C(4,1),CF(4,1),T(2),R(2,20)
COMMON /FLOCK6/ AMEAN(21),STDEV(21),VLEW(21,25)
COMMON /FLOCK7/ AMEAN(40),VALL(40),VALU(40),KDIM,YMIN,CY
DIMENSION NV(10),NVC(10),NCC(25)
DIMENSION VMAXC(25)
DATA NRC,NV,NVC,NCC /1,1,1/
C READ NUMBER OF DATA POINTS PER DISTRIBUTION
READ(5,1001) N
C READ NUMBER OF ITERATIONS
READ(5,1002) ITNS
PRINT 1003,ITNS
C READ NUMBER OF CONTROL VARIABLES, NV, NUMBER OF CONTROL
C VARIABLE CATEGORIES, NC, AND MAXIMUM NUMBER OF CONTROL VARIABLES
C IN A CATEGORY, MAXNO
READ(5,1004) NV,NC,MAXNO
PRINT 2001,NV,NC,MAXNO
C READ NUMBER OF CONTROL VARIABLES IN EACH CATEGORY.
READ(5,1005) (NVC(I), I=1,NC)
C READ NUMBER OF LINEAR CONSTRAINTS
READ(5,1006) NCCNS
PRINT 2002,NCCNS
C READ VMIN, VMAX, AND VINT FOR EACH CONTROL VARIABLE.
DO 10 J=1,NV
READ(5,1004) VMIN(J), VMAX(J), VINT(J)
PRINT 2004,J,VMIN(J),VMAX(J),VINT(J)
VMAXC(J)=VMAX(J)
10 CONTINUE
C CLEAR A MATRIX.
DO 10 J=1,NCCNS
DO 10 I=1,NV
A(I,J)=0.
10 CONTINUE
C READ A COEFFICIENTS.
READ(5,1003) I,J,A(I,J),LAST
PRINT 2003,I,J,A(I,J)
IF (LAST.EQ.0) GO TO 15
C READ CONSTRAINT LEVELS.
DO 20 I=1,NCCNS
READ(5,1005) F(I)
F(I)=F(I)
20 PRINT 2005,I,F(I)
C READ STOCHASTIC EXOGENOUS SYSTEM INPUT DATA
DO 25 I=1,NV
READ(5,1004) (CPC(J,I),J=1,10)
READ(5,1004) (EYC(J,I),J=1,10)
READ(5,1004) (CFS(J,I),J=1,10)
READ(5,1004) (EYS(J,I),J=1,10)
1005 FORMAT(10F8,2)
READ(5,1004) F(1,1),F(2,1)
READ(5,1004) (FHS(J,1),J=1,11)
DO 25 J=1,14
READ(5,1004) Y(J,1)
25 CONTINUE
C READ PREFERENCE DATA.
READ(5,500) YMIN,CY,KDIM
500 FORMAT(2F10,2,10)
DO 30 MPM=1,KDIM
READ(5,501) ARG(MPM),VALL(MPM),VALU(MPM)
501 FORMAT(1H,10,2,5V,2E20,13)
C ITERATION INITIALIZATION
DO 100 IT=1,ITNS
C SET ALL CONTROL VARIABLES AND THEIR INDICATOR VARIABLES TO 0 AND
C RESET VMAX
DO 35 J=1,NV
IND(J)=0
VMAX(J)=VMAXC(J)
35 V(J)=0
C RESET CONSTRAINT LEVELS
DO 40 I=1,NCCNS
F(I)=F(I)
40 CONTINUE
C RESET MIN,NEX,ILO,IMI
MIN=0
NEX=0
ILO=1
IMI=0
C BEGIN TO GENERATE PLAN
DO 65 I=1,NC
ILO=IMI+1
IMI=IMI+NVC(I)
IF (ILO.EQ.0) GO TO 55
50 CALL SELECT(ILO,IMI,J)
IF (IND(J).NE.0) GO TO 50
NEX(I)=NEX(I)+1
IND(J)=NEX(I)
CALL LEVCET(J)
CALL CHECK(J,NCCNS,NIN)
IF (NEX(I).LT.NVC(I)) GO TO 50
IF (NIN.GE.MAXNO) GO TO 70
GO TO 65
55 CONTINUE
DO 60 M=ILO,IMI
CALL LEVCET(M)
60 CONTINUE
65 CONTINUE
70 CALL EXPAND(F(I))

```

```

      CALL DISGEN
      CALL APTFC(NV)
      CONTINUE
100  DO 109 J=1,NF
      PRINT 2014,J,APEAN(J),STD(J)
      DO 110 J=1,NV
110  PRINT 2007,J,J,VLFV(J,J)
      DO 115 KK=1,N0
115  PRINT 2011,F(J,KK)
120  CONTINUE
1200 FORMAT(1P)
2000 FORMAT(5X,*,NUMBER OF ITERATIONS*,I8)
1001 FORMAT(3F5)
2001 FORMAT(5X,*,V=*,I5,*, NC=*,I5,*, MAXNO=*,I5)
1002 FORMAT(5F10)
2002 FORMAT(5X,*,NUMBER OF LINEAR CONSTRAINTS*,I8)
1003 FORMAT(210,F10.2,I5)
2003 FORMAT(5X,*,PO=*,I5,*, COLUMN=*,I5,*, A(I,J)=,F12.3)
1004 FORMAT(7F10.2)
2004 FORMAT(5X,*,CONTROL VARIABLE=*,I4,5X,*,VMIN=*,F8.2,5X,*,VMAX=*,F8.2,
15X,*,VINT=*,F8.2)
1005 FORMAT(F10.2)
2005 FORMAT(5X,*,CONSTRAINT ON RESOURCE=*,I5,F15.2)
2007 FORMAT(5X,*,LEVEL FOR CONTROL VARIABLE=*,I5,5X,F10.2)
2011 FORMAT(F15.3)
2014 FORMAT(1X,*,EFFICIENT PLAN=*,I3,3X,*, EXPECTED OUTCOME=*,F10.2,3X,*, STD=
1,F10.2)
1020 FORMAT(17Y,2F10.2)
1021 FORMAT(2Y5)
1022 FORMAT(1Y,11F10.2)
1023 FORMAT(10X,F10.2)
      END

```

```

      SUBROUTINE SELECT(M,N,J)
      DIFF=M+.9999
      R=DIFF*(-1)
      J=DIFF+M
      RETURN
      END

```

```

      SUBROUTINE LEVSET(I)
      COMMON /FLOCK1/ V(25),VMIN(25),VMAX(25),VINT(25)
      DIFF=(VMAX(I)-VMIN(I))/VINT(I)+.9999
      R=DIFF*(-1)
      L=DIFF+VMIN(I)/VINT(I)
      V(I)=FLOCAT(L)+VINT(I)
      RETURN
      END

```

```

      SUBROUTINE CHECK(J,NCONS,NIN)
      COMMON /FLOCK1/ V(25),VMIN(25),VMAX(25),VINT(25)
      COMMON /FLOCK2/ A(25,25),F(25),INC(25)
      IF(V(J).GT.0.0) GO TO 25
      DO 10 I=1,NCONS
      IF(A(I,J).EQ.0.0) GO TO 10
      IF(F(I).GE.(A(I,J)+V(J))) GO TO 10
      LEV=I/A(I,J)
      V(J)=LEV
      IF(V(J).LT.VMIN(J)) V(J)=0.0
      IF(V(J).GT.VMAX(J)) V(J)=VMAX(J)
10  CONTINUE
      IF(V(J).EQ.0.0) GO TO 25
      NIN=NIN+1
      DO 15 I=1,NCONS
15  F(I)=F(I)-V(J)*A(I,J)
      IF(J.NE.1) GO TO 25
      DO 20 I=0,3
20  VMAX(I)=VMAX(I)+V(J)
25  RETURN
      END

```

```

      SUBROUTINE EXPAND(ACRES)
      COMMON /FLOCK1/ V(25),VMIN(25),VMAX(25),VINT(25)
      COMMON /FLOCK2/ A(25,25),F(25),INC(25)
      IF(F(1).LE.0.0) GO TO 25
      IF(V(2).EQ.0.0) GO TO 25
      IF(V(3).EQ.0.0) GO TO 10
      IF(INC(3).LT.INC(2)) GO TO 5
      V(2)=ACRES+V(1)-V(3)
      RETURN
5  V(3)=ACRES+V(1)-V(2)
      RETURN
10  V(2)=ACRES+V(1)
      RETURN
15  V(3)=ACRES+V(1)
20  RETURN
      END

```

```

SUBROUTINE NSTOO(NV)
COMMON /FLOCK1/ V(25),VMIN(25),VPAX(25),VINT(25)
COMMON /FLOCK5/ NE,NE,NG,NP,C(41),CP(41),T(20),R(21,20)
COMMON /FLOCK6/ AMEAN(21),STD(21),VLEV(21,25)
DIMENSION CE1(21,10),CE2(21,10),LFLAG(21)
IFLAG=1
DO 100 I=1,21
  LFLAG(I)=0
  DO 101 J=1,NV
    VLEV(J,I)=V(I)
    VAS=0.0
    AMEAN(NP)=0.0
    PNE=0.0
    DO 102 J=1,NP
      AMEAN(NP)=AMEAN(NP)+(1.0/PNE)*R(NP,J)
      VAS=VAS+(1.0/PNE)*R(NP,J)*R(NP,J)
      T(J)=R(NP,J)
    CONTINUE
    VAS=VAS-AMEAN(NP)*AMEAN(NP)
    STD(NP)=SQRT(VAS)
    CALL RORBER
    DO 102 J=1,NP
      R(NP,J)=T(J)
    CONTINUE
    DO 103 I=1,NG
      B1=0.0
      B2=0.0
      DO 105 K=1,NP
        B1=B1+(1.0/END)*U1(R(NP,K),I)
        B2=B2+(1.0/END)*U2(R(NP,K),I)
      CONTINUE
      CE1(I,I)=U1(B1,I)
      CE2(I,I)=U2(B2,I)
    CONTINUE
    IF(NP.EQ.1) GO TO 115
    I=NP
    DO 106 J=1,NE
      K=1
      IF(CE1(I,K).GE.CE1(J,K).AND.CE2(I,K).GE.CE2(J,K)) GO TO 60
      IF(CE1(I,K).LE.CE1(J,K).AND.CE2(I,K).LE.CE2(J,K)) GO TO 61
      NSD=0
      GO TO 107
    60 CALL CUMCAL(J,I)
      CALL SD(CEU,K)
      IF(CEU(1,K).EQ.0) GO TO 62
      IF(CE1(I,K).EQ.CE1(J,K).AND.CE2(I,K).EQ.CE2(J,K)) GO TO 61
      NSD=1
      GO TO 107
    62 NSD=1
      GO TO 107
    61 CALL CUMCAL(I,J)
      CALL SD(CEU,K)
      IF(CEU(1,K).EQ.0) GO TO 63
      NSD=1
      GO TO 107
    63 NSD=1
      RETURN
    CONTINUE
    IF(NSD.EQ.0) GO TO 106
    IF(IFLAG.EQ.0) GO TO 120
    CE1(J,K)=CE1(I,K)
    CE2(J,K)=CE2(I,K)
    STD(J)=STD(I)
    AMEAN(J)=AMEAN(I)
    DO 110 MM=1,NP
      R(J,MM)=R(I,MM)
    DO 111 MM=1,NV
      VLEV(J,MM)=V(MM)
    IFLAG=1
    GO TO 100
  120 LFLAG(IFLAG)=J
    IFLAG=IFLAG+1
  116 CONTINUE
    NE=NE-IFLAG+1
    IF(NE.GT.20) NE=20
  115 NP=NP+1
    IF(IFLAG.LE.1.OR.LFLAG(1).GT.NE) GO TO 170
    J1=LFLAG(1)+1
    J2=2
    LE=LFLAG(1)
    DO 125 KK=LE,NE
      DO 121 KA=1,21
        IF(J2.GT.(IFLAG-1)) GO TO 122
        IF(J1.LT.LFLAG(J2)) GO TO 122
        J1=J1-1
        J2=J2+1
      121 CE1(KK,K)=CE1(J1,K)
        CE2(KK,K)=CE2(J1,K)
        STD(KK)=STD(J1)
        AMEAN(KK)=AMEAN(J1)
        DO 123 MM=1,NP
          R(KK,MM)=R(J1,MM)
        DO 124 MM=1,NV
          VLEV(KK,MM)=VLEV(J1,MM)
        J1=J1+1
      125 CONTINUE
    RETURN
  130 END

```

```

FUNCTION U1(X,JJ)
COMMON /FLOCK5/ ND,NE,NG,NP,C(41),CP(41),T(20),R(21,20)
COMMON /FLOCK6/ ARG(400),VALL(400),VALU(400),KDIM,YMIN,CV
DIMENSION X
N=KI*IF-1
U1=TAPEX(VALL,YMIN,CV,X,DUMMY)
RETURN
END

```

```

FUNCTION UT2(X,JJ)
COMMON /PLOCK45/ ND,NE,NG,NP,C(41),CP(41),T(20),R(21,20)
COMMON /PLOCKK/ ARG(400),VALL(400),VALU(400),KDIM,YMIN,DY
DUMMY=X
K=KDIM-1
UT2=TAFEXE(VALU,YMIN,DY,K,DUMMY)
RETURN
END

```

```

FUNCTION UT1(X,JJ)
COMMON /PLOCK45/ ND,NE,NG,NP,C(41),CP(41),T(20),R(21,20)
COMMON /PLOCKK/ ARG(400),VALL(400),VALU(400),KDIM,YMIN,DY
K=KDIM
UT1=TAFEX(ARG,VALL,X,K)
RETURN
END

```

```

FUNCTION UT2(X,JJ)
COMMON /PLOCK45/ ND,NE,NG,NP,C(41),CP(41),T(20),R(21,20)
COMMON /PLOCKK/ ARG(400),VALL(400),VALU(400),KDIM,YMIN,DY
K=KDIM
UT2=TAFEX(ARG,VALL,X,K)
RETURN
END

```

```

FUNCTION TAFEXE(VAL,SMALL,DIFF,K,DUMMY)
DIMENSION VAL(1)
DUM=DUMMY-SMALL
T=MAX(1.0,DUM/DIFF,1.0),K)
TAFEXE=(VAL(I+1)-VAL(I))*((DUM-FLOAT(I-1)*DIFF)/DIFF+VAL(I)
RETURN
END

```

```

FUNCTION TAFEX(VAL,ARG,DUMMY,K)
DIMENSION VAL(1),ARG(1)
DO 1 J=2,K
IF(DUMMY.GT.ARG(J)) GO TO 1
TAFEX=(DUMMY-ARG(J-1))*(VAL(J)-VAL(J-1))/
1 (ARG(J)-ARG(J-1))*VAL(J-1)
RETURN
1 CONTINUE
J=K
GO TO 2
END

```

```

SUBROUTINE RORDER
COMMON /PLOCK5/ ND,NE,NG,NP,C(41),CP(41),T(20),R(21,20)
ND1=ND-1
DO 400 I=1,ND1
AM=T(I)
N=I
II=I+1
DO 401 J=II,ND
IF(AM.LE.T(J)) GO TO 401
N=J
AM=T(J)
401 CONTINUE
T(I)=T(I)
T(I)=AM
400 CONTINUE
RETURN
END

```

```

SUBROUTINE CUMCAL(M1,M2)
COMMON /PLOCK5/ ND,NE,NG,NP,C(41),CP(41),T(20),R(21,20)
K1=1
K2=1
C(I)=0.0
CP(I)=0.0
PND=ND
ND2=ND+2+1
DO 500 I=2,ND2
IF(K1.GT.ND.AND.K2.GT.ND) GO TO 550
IF(K1.GT.ND) GO TO 553
IF(K2.GT.ND) GO TO 551
IF(C(M1,K1)+C(M2,K2)) 551,552,553
551 C(I)=C(I-1)+(1.0/PND)
CP(I)=R(M1,K1)
K1=K1+1
GO TO 500
552 C(I)=C(I-1)
CP(I)=R(M1,K1)
K1=K1+1
K2=K2+1
GO TO 500
553 C(I)=C(I-1)-(1.0/PND)
CP(I)=R(M2,K2)
K2=K2+1
GO TO 500
550 C(I)=0.0
CP(I)=0.0
500 CONTINUE
RETURN
END

```

```

SUBROUTINE SD(MU,II)
COMMON /FLOCK/ NC,NE,NG,NP,C(41),CP(41),T(20),R(21,20)
EXTERNAL U1,U2
REAL MU
MU=0.0
NL=2*ND
ND=2*ND
DO 600 J=1,ND2
IF (NL.LE.1) GO TO 653
IF (C(NL)) 651,652,653
652 NL=NL-1
MU=0.0
GO TO 600
653 MU=MU+U2(CP(NL+1),II)-U2(CP(NL),II))*C(NL)
NL=NL-1
DO 601 J=1,ND2
IF (NL.LE.1) GO TO 656
TU=(U2(CP(NL+1),II)-U2(CP(NL),II))*C(NL)
IF (MU+TU) 654,652,655
655 MU=MU+TU
NL=NL-1
601 CONTINUE
654 TI=MU/C(NL)+U2(CP(NL+1),II)
C=U12(TI,II)
MU=0.0
CP(NL+1)=Q
GO TO 600
651 MU=MU+U1(CP(NL+1),II)-U1(CP(NL),II))*C(NL)
NL=NL-1
DO 602 K=1,ND2
IF (NL.LE.1) GO TO 657
TU=(U1(CP(NL+1),II)-U1(CP(NL),II))*C(NL)
IF (MU+TU) 656,652,657
656 MU=MU+TU
NL=NL-1
602 CONTINUE
657 TS=MU/C(NL)+U1(CP(NL+1),II)
G=U11(TS,II)
MU=0.0
CP(NL+1)=G
CONTINUE
RETURN
END

```

[illegible]

```

      ZA=(CPS(J,KDIM)-FXS(J,KDIM))/EXS(J,KDIM)
      ZP=CPS
      ZC=RASH/DASF+V(12)-FSC/FS
      DSC=ES*(V(9)+7A+V(12)+7F+V(11)+ZC)
      IF(DSC.LE.0.0) GO TO 8
      IF((FSC+SC).GT.SC+X) DSC=SC+X-BSC
      RSC=FSC+DSC+CPS(J,KDIM)
      ESC=FSC+DESC
      VM=J+1
8     DCPD=(CPC(M,KDIM)-CPC(J,KDIM))/(F(M)-F(J))
      DCPD=(CPC(M,KDIM)-CPC(J,KDIM))/(F(M)-F(J))
9     CONTINUE
      DO 50 J=1,M
      IF(J.GT.4) GO TO 43
      RH=PC(J,KDIS)
      IF(RCMH.LE.0.0.OR.J.GT.4) GO TO 15
      IF(RCPH.GT.RH) GO TO 10
      R(J)=RACP
      RCPH=R(J)
      RH=RH-G(J)*.757
      GO TO 15
10     G(J)=RH/.757
      RCPH=RCPH-RH
      RACP=RACP-G(J)
      IF(RCAT(J).LT.V(4)) GO TO 40
      RSPH=RSPH+RCPH
      RASH=RASH+RACP
      RCPH=0.0
      RACP=0.0
      DACP=DACP-RACP
      EC=100.0*LACP
      CCMAX=1.5*EC
      DASF=DASF+RACP
      ES=77.0*DASF
      SCHAY=1.5*ES
      GO TO 40
15     IF(RCPH.LE.0.0) GO TO 40
      IF(J.LT.3) GO TO 40
      IF(RSPH.GT.RH) GO TO 20
      S(J)=RACP
      RSPH=0.0
      GO TO 40
20     S(J)=RH/.757
      RSPH=RSPH-RH
      RASH=RASH-S(J)
      RACP=RACP-G(J)
40     RASH=RASH-S(J)
43     VM=J+4
      IF(V(2).EQ.0.0) GO TO 41
      ZA=(CPC(M,KDIM)-EXC(M,KDIM))/EXC(M,KDIM)
      ZC=EXC
      ZC=RASH/DASF+V(12)-PCC/EC
      DCC=EC*(V(9)+7A+V(12)+7F+V(7)+ZC)
      IF(DCC.LE.0.0) GO TO 41
      IF((FCC+CC).GT.CC+X) DCC=CC+X-PCC
      RCC=FCC+DCC+CPC(M,KDIM)
      DCC=FCC+DCC
41     IF(V(3).EQ.0.0) GO TO 42
      ZA=(CPS(M,KDIM)-EXS(M,KDIM))/EXS(M,KDIM)
      ZC=CPS
      ZC=RASH/DASF+V(12)-FSC/FS
      DSC=ES*(V(9)+7A+V(12)+7F+V(11)+ZC)
      IF(DSC.LE.0.0) GO TO 42
      IF((FSC+SC).GT.SCHAY) DESC=SC+X-BSC
      RSC=FSC+DSC+CPS(M,KDIM)
      ESC=FSC+DESC
      VM=M+1
42     IF(M.GT.11) GO TO 57
      DCPD=(CPC(M,KDIS)-CPC(M,KDIS))/(F(MM)-F(M))
      DCPD=(CPC(MM,KDIS)-CPC(M,KDIS))/(F(MM)-F(M))
50     CONTINUE
60     DO 60 J=7,11
      RH=PC(J,KDIS)
      IF(RASH.LE.0.0.OR.J.EQ.11) GO TO 70
      ASH=RH/.502
      DO 65 KK=3,6
      JS=4*KK+J
      IF(ASH-S(KK)) 61,61,62
61     XLEV(JS)=ASH
      RASH=RASH-ASH
      S(KK)=S(KK)-ASH
      GO TO 60
62     XLEV(JS)=S(KK)
      RASH=RASH-S(KK)
      RH=RH-S(KK)*.502
      ASH=ASH-S(KK)
      S(KK)=0.0
      IF(RASH.LE.0.0) GO TO 90
      IF(RASH.LE.0.0) GO TO 70
65     CONTINUE
70     IF(RASH.GT.0.0.AND.J.LT.11) GO TO 90
      ICM=RH/.418
      DO 76 KK=1,4
      IF(KK-2) 71,71,72
71     JS=5*KK+J-11
      GO TO 73
72     IF(J.EQ.7) GO TO 90
      JS=4*KK+J-9
73     IF(ACH-G(KK)) 74,74,75
74     XLEV(JS)=ACH
      RACH=RACH-ACH
      G(KK)=G(KK)-ACH
      GO TO 70
75     XLEV(JS)=G(KK)
      RACH=RACH-G(KK)
      RH=RH-G(KK)*.418
      ACH=ACH-G(KK)

```

```

      G(KK)=C.*
      IF (RACH.LF.0.0) GO TO 91
75  CONTINUE
85  CONTINUE
91  IF (RACH.LT.0.0) RACH=0.0
      IF (RASH.LT.0.0) RASH=0.0
      R(NP,KDIS)=R(NP,KDIS)-4.*RASH
      R(NP,KDIS)=R(NP,KDIS)-(1-HC(14)*0.5*Y(1P,KDIS)*AC(1))*RACH
      RCH=CH+0.7*Y(1P,KDIS)*RACH
      DO 92 J=1,10
      RCH=CH+XLEV(J)*Y(J,KDIS)
      R(NP,KDIS)=R(NP,KDIS)-(1-HC(J)*Y(J,KDIS)*AC(1))*XLEV(J)
92  CONTINUE
      DO 93 J=10,14
      RCH=CH+XLEV(J)*Y(J,KDIS)
      R(NP,KDIS)=R(NP,KDIS)-(1-HC(J)*Y(J,KDIS)*AC(2))*XLEV(J)
93  CONTINUE
      R(NP,KDIS)=R(NP,KDIS)+RCC*P(1,KDIS)+(BCH-BC(1)*RSC*P(2,KDIS)+(BSH-
100  BS(1)-BS(2)*RSC)*P(3,KDIS)
      CONTINUE
      RETURN
      END

```


BIBLIOGRAPHY

BIBLIOGRAPHY

- Anderson, J. R., 1974a. "Simulation: Methodology and Application in Agricultural Economics." Review of Marketing and Agricultural Economics 42:3-55.
- _____. 1974b. "Sparse Data, Estimational Reliability, and Risk-Efficient Decisions." American Journal of Agricultural Economics 56:564-72.
- _____. 1975. "Programming for Efficient Planning Against Non-Normal Risk." Australian Journal of Agricultural Economics 19:94-107.
- _____. 1976. "Risk Efficient Monte Carlo Programming: A Manual." Miscellaneous Publication No. 6, Department of Agricultural Economics and Business Management, University of New England, Armidale, Australia.
- Anderson, J. R., J. L. Dillon, and J. B. Hardaker, 1977. Agricultural Decision Analysis. Ames, Iowa: The Iowa State University Press.
- Anderson, J. R. and H. E. Doran, 1978. "Some Properties of Simple Functions of Random Variables." Review of Marketing and Agricultural Economics 46:37-47.
- Anderson, T. W., 1958. An Introduction to Multivariate Statistical Analysis. New York: John Wiley and Sons, Inc.
- Aoki, Masanao, 1967. Optimization of Stochastic Systems. New York: Academic Press, Inc.
- _____. 1975. "Output Decisions by a Firm: An Example of a Dual Control Problem with Information Externalities," in Adaptive Economic Models, Day, R. H. and T. Groves, eds. New York: Academic Press, pp. 157-194.
- Arrow, K. J., 1971. Essays in the Theory of Risk-Bearing. Chicago: Markham Publishing Company.
- Arrow, K. J. and M. Kurz, 1970. Public Investment, the Rate of Return, and Optimal Fiscal Policy. Baltimore: Johns Hopkins Press.

- Barnett, V., 1975. "Probability Plotting Methods and Order Statistics." Applied Statistics 24:95-108.
- Barry, P. J. and D. R. Willman, 1976. "A Risk Programming Analysis of Forward Contracting with Credit Constraints." American Journal of Agricultural Economics 58:62-70.
- Black, J. R., W. A. Tinsley, E. Fuller, and C. Cuykendall, no date. "Corn, Soybean Farm Planning Guide." Telplan Program 18, Form 1, Department of Agricultural Economics, Michigan State University, East Lansing, Michigan.
- Boussard, J. M. and M. Petit, 1967. "Representation of Farmers' Behavior under Uncertainty with a Focus-Loss Constraint." Journal of Farm Economics 49:869-80.
- Bradford, L. A. and G. L. Johnson, 1953. Farm Management Analysis. New York: John Wiley and Sons, Inc.
- Chapman, L. J., 1967. "Illusory Correlation in Observational Report." Journal of Verbal Learning and Verbal Behavior 6:151-155.
- Charnes, A. and W. W. Cooper, 1959. "Chance Constrained Programming." Management Science 6:73-79.
- Chow, G. C., 1973. "Effect of Uncertainty on Optimal Control Problems." International Economic Review 14:632-645.
- Churchman, C. W., 1968. The Systems Approach. New York: Dell Publishing, Inc.
- Cocks, K. D., 1968. "Discrete Stochastic Programming." Management Science 15:72-79.
- Coleman, D. R. and A. L. Saipé, 1976. "Modeling Bivariate Relationships for Simulation," unpublished manuscript.
- Coleman, D. R. and A. L. Saipé, 1977. "Simulating a Lognormal Time Series with Prescribed Serial Correlation." Management Science 23:1363-1364.
- Conklin, F. S. A. E. Baquet, and A. N. Halter, 1977. "A Bayesian Simulation Approach for Estimating Value of Information: An Application to Frost Forecasting." Technical Bulletin 136, Agricultural Experiment Station, Oregon State University, Corvallis, Oregon.
- Cyert, R. M. and J. G. March, 1963. A Behavioral Theory of the Firm. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

- Day, R. H., 1963. Recursive Programming and Production Response. Amsterdam: North Holland Publishing Company.
- _____. 1964. "A Behavioral Theory of the Firm by Cyert and March." Econometrica 32:461-464.
- _____. 1971. "Rational Choice and Economic Behavior." Theory and Decision 1:229-251.
- _____. 1975. "Adaptive Processes in Economic Theory," in Adaptive Economic Models, Day, R. H. and T. Groves, eds. New York: Academic Press, pp. 1-38.
- Dent, J. B. and J. R. Anderson, eds., 1971. Systems Analysis in Agricultural Management. Sydney, Australia: John Wiley and Sons, Inc.
- Dent, J. B. and S. C. Thompson, 1968. "The Application of Monte Carlo Techniques to the Feedmix Problem." Farm Economist 11:230-248.
- Donaldson, G. F. and J. P. G. Webster, 1968. "An Operating Procedure for Simulation Farm Planning--Monte Carlo Method." Department of Agricultural Economics, Wye College, University of London.
- Dreyfus, S. E., 1968. "Introduction to Stochastic Optimization and Control," in Stochastic Optimization and Control, Karreman, H. F., ed. New York: John Wiley and Sons, Inc., pp. 3-24.
- Ehrlich, I. and G. S. Becker, 1972. "Market Insurance, Self-Insurance, and Self-Protection." Journal of Political Economy 80:627-648.
- Eisgruber, L. M. and G. E. Lee, 1971. "A System Approach to Studying the Growth of the Farm Firm," in Systems Analysis in Agricultural Management, Dent, J. B. and J. R. Anderson, eds. Sydney, Australia: Wiley and Sons, Inc.
- Fishburn, P. C., 1970. Utility Theory for Decision Making. New York: John Wiley and Sons, Inc.
- Forrester, J. W., 1961. Industrial Dynamics. New York: John Wiley and Sons, Inc.
- Freund, R. J., 1956. "The Introduction of Risk into a Programming Model." Econometrica 24:253-63.
- Friedman, M. and L. J. Savage, 1948. "Utility Analysis of Choices Involving Risk." Journal of Political Economy 56:279-304.
- Grayson, C. J., 1960. Decisions Under Uncertainty: Drilling Decisions by Oil and Gas Operators. Boston: Graduate School of Business Administration, Harvard University.

- Hadar, J. and W. R. Russell, 1969. "Rules for Ordering Uncertain Prospects." American Economic Review 59:25-34.
- Hanoch, G. and H. Levy, 1969. "The Efficiency Analysis of Choices Involving Risk." Review of Economic Studies 36:335-46.
- Hazell, P. B. R., 1970. "Game Theory--An Extension of its Application to Farm Planning under Uncertainty." Journal of Agricultural Economics 21:239-52.
- _____. 1971. "A Linear Alternative to Quadratic and Semi-variance Programming for Farm Planning under Uncertainty." American Journal of Agricultural Economics 53:53-62.
- Heidhues, T., 1966. "A Recursive Programming Model of Farm Growth in Northern Germany." Journal of Farm Economics 48:668-684.
- Heifner, R. L., 1972. "Optimal Hedging Levels and Hedging Effectiveness in Cattle Feeding." Agricultural Economics Research 24:25-36.
- Hicks, J. R., 1946. Value and Capital, 2nd edition. London: Oxford University Press.
- Hirshleifer, J., 1970. Investment, Interest and Capital. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.
- Hogarth, R. M., 1975. "Cognitive Processes and the Assessment of Subjective Probability Distributions." Journal of the American Statistical Association 70:271-91.
- Holt, C. C., F. Modigliani, J. F. Muth, and H. A. Simon, 1960. Planning Production, Inventories and Work Force. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.
- Howard, R. A., J. E. Matheson, and D. W. North, 1972. "The Decision to Seed Hurricanes." Science 176:1191-202.
- Johnson, G. L., 1961a. "Some Reflections on the Nature of Managerial Problems," in Managerial Processes of Midwestern Farmers, Johnson, G. L., A. N. Halter, H. R. Jensen, and D. W. Thomas, eds. Ames, Iowa: The Iowa State University Press, pp. 142-149.
- _____. 1961b. "Implications of the IMS for Study of Responses to Price," in Managerial Processes of Midwestern Farmers, Johnson, G. L., A. N. Halter, H. R. Jensen, and D. W. Thomas, eds. Ames, Iowa: The Iowa State University Press, pp. 150-169.
- _____. 1976. "Philosophic Foundations: Problems, Knowledge, and Solutions." European Review of Agricultural Economics 3:207-234.

Jo

J

J

V

- Johnson, G. L. and C. F. Lard, 1961. "Knowledge Situations," in Managerial Processes of Midwestern Farmers, Johnson, G. L., A. N. Halter, H. R. Jensen, and D. W. Thomas, eds. Ames, Iowa: The Iowa State University Press, pp. 41-54.
- Johnson, G. L. and C. L. Quance, eds., 1972. The Overproduction Trap in U.S. Agriculture. Baltimore, Johns Hopkins Press.
- Johnson, G. L. and L. K. Zerby, 1973. What Economists Do About Values. East Lansing, Michigan: Department of Agricultural Economics, Center for Manpower and Public Affairs, Michigan State University.
- Kahneman, D. and A. Tversky, 1973. "On the Psychology of Prediction." Psychological Review 80:237-251.
- Kahneman, D. and A. Tversky, 1972. "Subjective Probability: A Judgement of Representativeness." Cognitive Psychology 3:430-454.
- Karreman, H. F., ed., 1968. Stochastic Optimization and Control. New York: John Wiley and Sons, Inc.
- Keeney, R. L., 1973. "A Decision Analysis with Multiple Objectives: The Mexico City Airport." Bell Journal of Economics and Management Science 4:101-17.
- Keeney, R. L. and H. Raiffa, 1976. Decisions with Multiple Objectives: Preferences and Tradeoffs. New York: Wiley and Sons, Inc.
- Kihlstrom, R. E. and L. J. Mirman, 1974. "Risk Aversion with Many Commodities." Journal of Economic Theory 8:361-388.
- Kirk, D. E., 1970. Optimal Control Theory. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.
- Kleindorfer, P. and H. Kunreuther, 1978. "Stochastic Horizons for the Aggregate Planning Problem." Management Science 24:485-497.
- Knight, F. H., 1921. Risk, Uncertainty and Profit. Boston: Houghton-Mifflin.
- Levy, H., 1973. "Stochastic Dominance, Efficiency Criteria and Efficient Portfolios: The Multi-Period Case." American Economic Review 63:986-994.
- Levy, H. and J. Paroush, 1974a. "Toward Multivariate Efficiency Criteria." Journal of Economic Theory 7:129-42.
- Levy, H. and J. Paroush, 1974b. "Multi-Period Stochastic Dominance." Management Science 24:428-441.

- Lewis, C. I., 1955. The Ground and Nature of the Right. New York: Columbia University Press.
- Lin, W. W. and H. S. Chang, 1978. "Specification of Bernoullian Utility Function in Decision Analysis." Agricultural Economics Research 30:30-36.
- Llewellyn, R. W., 1965. "FORDYN--An Industrial Dynamic Simulator." Department of Industrial Engineering, North Carolina State University, Raleigh, North Carolina.
- Low, A. R. C., 1974. "Decision Taking under Uncertainty: A Linear Programming Model of Peasant Farmer Behavior." Journal of Agricultural Economics 25:311-22.
- McInerney, J. P., 1969. "Linear Programming and Game Theory Models." Journal of Agricultural Economics 20:269-278.
- McKinnon, R. I., 1967. "Futures Markets, Buffer Stocks, and Income Stability for Primary Producers." Journal of Political Economy 75:844-861.
- MacRae, E. C., 1975. "An Adaptive Learning Rule for Multiperiod Decision Problems." Econometrica 43:893-906.
- Manetsch, T. J., 1978. "Theory and Practice of Model Building and Simulation," in Agricultural Sector Planning, Rossmiller, G. E., ed. Department of Agricultural Economics, Michigan State University, East Lansing, Michigan.
- Manetsch, T. J. and G. L. Park, 1977a. System Analysis and Simulation with Applications to Economic and Social Systems, Part I, 3rd edition. Department of Electrical Engineering and System Science, Michigan State University, East Lansing, Michigan.
- Manetsch, T. J. and G. L. Park, 1977b. System Analysis and Simulation with Applications to Economic and Social Systems, Part II, 3rd edition. Department of Electrical Engineering and System Science, Michigan State University, East Lansing, Michigan.
- Markowitz, H., 1959. Portfolio Selection--Efficient Diversification of Investments. New York: Wiley and Sons.
- Massé, P., 1962. Optimal Investment Decisions. Englewood Cliffs, New Jersey: Prentice, Inc.
- Meyer, J., 1977a. "Choice Among Distributions." Journal of Economic Theory 14:326-36.

- Meyer, J., 1977b. "Second Degree Stochastic Dominance with Respect to a Function." International Economic Review 18:477-87.
- _____. 1977c. "Further Applications of Stochastic Dominance to Mutual Fund Performance." Journal of Financial and Quantitative Analysis 12:235-242.
- Modigliani, F. and K. J. Cohen, 1961. The Role of Anticipations and Plans in Economic Behavior and Their Use in Economic Analysis and Forecasting. Studies in Business Expectations and Planning, No. 4, Bureau of Economic and Business Research, University of Illinois, Urbana, Illinois.
- Mood, A. M. and F. A. Graybill, 1963. Introduction to the Theory of Statistics. New York: McGraw-Hill, Inc.
- Naylor, T. H., J. L. Balintty, D. S. Burdick, K. Chu, 1966. Computer Simulation Techniques. New York: John Wiley and Sons, Inc.
- Newman, T. G. and P. L. Odell, 1971. The Generation of Random Variates. New York: Hafner Publishing Company.
- Nott, S., G. Schwab, S. Harsh, J. Black, M. Kelsey, 1977. "Enterprise Budgets Michigan, 1977." Agricultural Economics Report No. 314, Department of Agricultural Economics, Michigan State University, East Lansing, Michigan.
- Officer, R. R. and A. N. Halter, 1968. "Utility Analysis in a Practical Setting." American Journal of Agricultural Economics 50:257-77.
- Phillips, D. T. and C. S. Beightler, 1972. "Procedure for Generating Gamma Variates with Non-Integer Parameter Sets." Journal of Statistical Computation and Simulation 1:203.
- Pratt, J. W., 1964. "Risk Aversion in the Small and in the Large." Econometrica 32:122-36.
- Pratt, J. W., H. Raiffa, and R. Schlaifer, 1965. Introduction to Statistical Decision Theory. New York: McGraw-Hill, Inc.
- Rae, A. N., 1971. "Stochastic Programming, Utility, and Sequential Decision Problems in Farm Management." American Journal of Agricultural Economics 53:448-60.
- Ramsey, F. P., 1931. "Truth and Probability," in The Foundations of Mathematics and Other Logical Essays, Braithwaite, R. B., ed. London: Routledge and Kegan Paul.
- Robison, L. J. and R. P. King, 1978. "Specification of Micro Risk Models for Farm Management and Policy Research." Agricultural Economics Report No. 349, Department of Agricultural Economics, Michigan State University, East Lansing, Michigan.

- Sage, A. P., 1968. Optimum Systems Control. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.
- Samuelson, P. A., 1967. "General Proof That Diversification Pays." Journal of Financial and Quantitative Analysis 2:1-13.
- Savage, L. J., 1971. "Elicitation of Personal Probabilities and Expectations." Journal of the American Statistical Association 66:783-801.
- Schlaifer, R., 1959. Probability and Statistics for Business Decisions. New York: McGraw-Hill, Inc.
- Schmidt, J. W. and R. E. Taylor, 1970. Simulation and Analysis of Industrial Systems. Homewood, Illinois: Richard D. Irwin, Inc.
- Spetzler, C. S. and C-A. S. Stael von Holstein, 1975. "Probability Encoding in Decision Analysis." Management Science 22:340-58.
- Stael von Holstein, C-A. S., 1970. Assessment and Evaluation of Subjective Probability Distributions. Stockholm: The Economic Research Institute at the Stockholm School of Economics.
- Tobin, J., 1958. "Liquidity Preference as Behavior Towards Risk." Review of Economic Studies 25:65-78.
- Tsiang, S. C., 1972. "The Rationale of the Mean-Variance Analysis, Skewness Preference, and the Demand for Money." American Economic Review 62:354-372.
- Tversky, A. and D. Kahneman, 1973. "Availability: A Heuristic for Judging Frequency and Probability." Cognitive Psychology 5:207-232.
- von Neumann, J. and O. Morgenstern, 1944. Theory of Games in Economic Behavior. Princeton, New Jersey: Princeton University Press.
- Wald, A., 1947. Sequential Analysis. New York: John Wiley and Sons, Inc.
- Ward, R. W. and L. B. Fletcher, 1971. "From Hedging to Pure Speculation: A Micro Model of Optimal Futures and Cash Market Positions." American Journal of Agricultural Economics 53:71-78.
- Watson, S. R., J. R. Weiss, and M. L. Donnell, 1979. "Fuzzy Decision Analysis." IEEE Transactions on Systems, Man, and Cybernetics SMC-9:1-9.
- Whitmore, B. A., 1970. "Third-Degree Stochastic Dominance." American Economic Review 60:457-459.

7

Winkler, R. L., 1967. "The Assessment of Prior Distributions in Bayesian Analysis." Journal of the American Statistical Association 62:776-800.

_____, 1969. "Scoring Rules and the Evaluation of Probability Assessors." Journal of the American Statistical Association 64:1073-1078.

Zadeh, L. A., 1973. "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes." IEEE Transactions on Systems, Man, and Cybernetics SMC-3:28-44.

Zellner, A., 1971. An Introduction to Bayesian Inference in Econometrics. New York: Wiley and Sons, Inc.

