

2.1

TWO REPRESENTATION THEOREMS
AND THEIR APPLICATION TO DECISION THEORY

by

CHEW, SOO HONG

M.A., Claremont Graduate School, 1977

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES
(Interdisciplinary Programme)
(Mathematics, Economics, Management Science)

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

November 1980

© Chew, Soo Hong, 1980

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study.

I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Interdisciplinary

The University of British Columbia
2075 Wesbrook Place
Vancouver, Canada
V6T 1W5

Date Feb. 23 1981

This dissertation consists of two parts. Part I contains the statements and proofs of two representation theorems. The first theorem, proved in Chapter 1, generalizes the quasilinear mean of Hardy, Littlewood and Polya by weakening their axiom of quasilinearity. Given two distributions with the same means, quasilinearity requires that mixtures of these distributions with another distribution in the same proportions share the same mean, regardless of the distribution that they are mixed with. We weaken the quasilinearity axiom by allowing the proportions that give rise to the same means to be different. This leads to a more general mean, denoted by $M_{\alpha\phi}$, which has the form:

$$M_{\alpha\phi}(F) = \phi^{-1}(\int_R \alpha\phi dF / \int_R \alpha dF),$$

where ϕ is continuous and strictly monotone, α is continuous and strictly positive (negative) and F is a probability distribution. The quasilinear mean, denoted by M_ϕ , results when the α function is constant. We showed, in addition, that the $M_{\alpha\phi}$ mean has the intermediate value property, and can be consistent with the stochastic dominance (including higher degree ones) partial order. We also generalized a well known inequality among quasilinear means, via the observation that the $M_{\alpha\phi}$ mean of a distribution F can be written as the quasilinear mean of a distribution F^α , where F^α is derived from F via α as the Radon-Nikodym derivative of F^α with respect to F .

We noted that the $M_{\alpha\phi}$ mean induces an ordering among probability distributions via the maximand, $\int_R \alpha\phi dF / \int_R \alpha dF$, that contains the (expected utility) maximand, $\int_R \phi dF$, of the quasilinear mean M_ϕ as a special case. Chapter 2 provides an alternative characterization of the above representation for simple probability measures on a more general outcome set

where mean values may not be defined. In this case, axioms are stated directly in terms of properties of the underlying ordering. We retained several standard properties of expected utility, namely weak order, solvability and monotonicity but relaxed the substitutability axiom of Pratt, Raiffa and Schlaifer, which is essentially a restatement of quasi-linearity in the context of an ordering.

Part II of the dissertation concerns one specific area of application - decision theory. Interpreting the $M_{\alpha\phi}(F)$ mean of Chapter 1 as the certainty equivalent of a monetary lottery F , the corresponding induced binary relation has the natural interpretation as 'strict preference' between lotteries. For non-monetary (finite) lotteries, we apply the representation theorem of Chapter 2. The hypothesis, that a choice agent's preference among lotteries can be represented by a pair of α and ϕ functions through the $M_{\alpha\phi}$ induced ordering, is referred to as alpha utility theory. This is logically equivalent to saying that the choice agent obeys either the mean value (certainty equivalent) axioms or the axioms on his strict preference binary relation.

Alpha utility theory is a generalization of expected utility theory in the sense that the expected utility representation is a special case of the alpha utility representation. The motivation for generalizing expected utility comes from difficulties it faced in the description of certain choice phenomena, especially the Allais paradox. These are summarized in Chapter 3.

Chapter 4 contains the formal statements of assumptions and the derivations of normative and descriptive implications of alpha utility theory. We stated conditions, taken from Chapter 1, for consistency with

stochastic dominance and global risk aversion and derived a generalized Arrow-Pratt index of local risk aversion. We also demonstrated how alpha utility theory can be consistent with those choice phenomena that contradict the implications of expected utility, without violating either stochastic dominance or local risk aversion. The chapter ended with a comparison of alpha utility with two other theories that have attracted attention; namely, Allais' theory and prospect theory.

Several other applications of the representation theorems of Part I are considered in the Conclusion of this dissertation. These include the use of the $M_{\alpha\phi}$ mean as a model of the equally-distributed-equivalent level of income (Atkinson, 1970), and as a measure of asymmetry of a distribution (Canning, 1934). The alpha utility representation can also be used to rank social situations in the sense of Harsanyi (1977). We ended by pointing out an open question regarding conditions for comparative risk aversion and stated an extension of Samuelson's (1967) conjecture that Arrow's impossibility theorem would hold if individuals and society express their preferences by von Neumann-Morgenstern utility functions.

INTRODUCTION 1

PART I TWO REPRESENTATION THEOREMS

1 GENERALIZING THE QUASILINEAR MEAN OF HARDY, LITTLEWOOD
AND POLYA 7

1.1 INTRODUCTION 7

1.2 AXIOMS OF MEAN VALUE 8

1.3 REPRESENTATION THEOREM 14

1.4 PROPERTIES OF THE $M_{\alpha\phi}$ MEAN 292 GENERALIZING THE EXPECTED UTILITY REPRESENTATION
THEOREM 36

2.1 INTRODUCTION 36

2.2 PRELIMINARY DEFINITIONS 36

2.3 AXIOMS 38

2.4 REPRESENTATION THEOREMS 40

PART II APPLICATION TO DECISION THEORY

BACKGROUND 59

3 CRITIQUE OF EXPECTED UTILITY THEORY 61

3.1 INTRODUCTION 61

3.2 SYSTEMATIC VIOLATION OF THE STRONG INDEPENDENCE
PRINCIPLE 64

| | | |
|-------|---|-----|
| 3.3 | CONCURRENCE OF RISK SEEKING AND RISK AVERTING BEHAVIOR | 70 |
| 3.4 | SOME PROBLEMS WITH PROBLEM REPRESENTATION | 77 |
| 3.5 | SUMMARY | 80 |
| 4 | A NEW THEORY | 81 |
| 4.1 | INTERPRETING MEAN VALUE AS CERTAINTY EQUIVALENT | 81 |
| 4.2 | REPRESENTATION OF A PREFERENCE BINARY RELATION | 86 |
| 4.3 | NORMATIVE IMPLICATIONS | 90 |
| 4.3.1 | Ratio Consistency | 91 |
| 4.3.2 | Assessment | 97 |
| 4.3.3 | Stochastic Dominance | 100 |
| 4.3.4 | Global Risk Aversion | 103 |
| 4.3.5 | Local Risk Aversion: The Arrow-Pratt Index | 105 |
| 4.4 | DESCRIPTIVE IMPLICATIONS | 108 |
| 4.4.1 | Systematic Violation of the Strong Independence Axiom | 108 |
| 4.4.2 | Stochastic Dominance | 114 |
| 4.4.3 | Local and Global Risk Properties: Concurrence of Risk Averting and Risk Seeking Behavior | 116 |
| 4.4.4 | Some Problems with Problem Representation | 122 |
| 4.5 | CRITIQUE OF ALLAIS' THEORY AND PROSPECT THEORY | 124 |
| 4.5.1 | Allais' Theory | 124 |
| 4.5.2 | Prospect Theory | 131 |
| 4.5.3 | Comparison | 137 |

CONCLUSION

5 CONCLUSION 141

5.1 SUMMARY 141

5.2 EXTENSIONS 143

TABLES

| | | |
|-----|---|-----|
| 3.1 | Summary of Empirical Results Relevant to the HIL0 Lottery Structure | 68 |
| 4.1 | Allowable Choice Patterns under Alpha Utility Theory | 110 |
| 4.2 | Comparison among Theories | 139 |

FIGURES

- 3.1 A "typical" von Neumann-Morgenstern utility function 62
- 3.2 Four decision problems 64
- 3.3 The composition of the $A_L(B_L)$ lottery from $A_0(B_0)$ 65
- 3.4 The HILO structure of three consequence lotteries 66
- 3.5 Standard lottery comparison 70
- 3.6 Examples of u_c and $u_{1/2}$ (based on data from Allais (1977), Appendix C) 74
- 3.7 Examples of $u_{3/4}$ and u_g (based on data from MacCrimmon et.al. (1972)) 75
- 3.8 Graphical representation of two lotteries 77
- 3.9 A sequential representation of lottery B 77
- 4.1 Ratio Consistency illustrated using barycentric coordinates 92
- 4.2 Geometric proof of the Ratio Consistency property 95
- 4.3 Probability Equivalent method 98
- 4.4 Test of Substitution axiom 98
- 4.5 Preference Pattern for $\alpha(I) < 1$ 109
- 4.6 Consistency Conditions for Stochastic Dominance 115
- 4.7 Conditions for Local Risk Aversion 117
- 4.8 An admissible alpha function 118
- 4.9 A pair of u_c and $u_{1/2}$ derived from an alpha utility decision maker 120
- 4.10 A pair of u_g and $u_{3/4}$ derived from an alpha utility decision maker 120
- 5.1 An alpha function that discriminates against the rich 145

PREFACE

The preface is perhaps a suitable place for an informal discussion of ideas leading to this dissertation. In their axiomatization of expected utility, von Neumann and Morgenstern likened the "utility", loosely speaking, of a lottery to the center of gravity of a mass distribution. In other words, the utility, $u(x_1, \dots, x_n; p_1, \dots, p_n)$, of a lottery, $(x_1, \dots, x_n; p_1, \dots, p_n)$, which pays x_i with probability p_i , should be the mean or average of the utilities associated with each probable outcome. This led to the expected utility expression:

$$u(x_1, \dots, x_n; p_1, \dots, p_n) = \sum_{i=1}^n p_i u(x_i).$$

Together with their elegant axiomatic characterization, there seemed to be a strong case for the adoption of expected utility. The Allais-Savage controversy however clouded the picture somewhat. Some, including Allais, suggested replacing the probabilities, $\{p_i\}_{i=1}^n$, with a more general set of weights, $\{\phi_i\}_{i=1}^n$, which depend on the lottery, and sum to unity. In this case,

$$u(x_1, \dots, x_n; p_1, \dots, p_n) = \sum_{i=1}^n \phi_i(x_1, \dots, x_n; p_1, \dots, p_n) u(x_i).$$

At this level of generality, the only testable implication is that the utility of a lottery is intermediate in value between the maximum and the minimum attainable utilities. This property, which may be called intermediate value property, is compatible with our intuition about the utility of a lottery as a mean value.

To impose more structure, one may consider restricting the ϕ_i weights to:

$$\phi_i(x_1, \dots, x_n; p_1, \dots, p_n) = \frac{w_i(x_1, \dots, x_n; p_1, \dots, p_n)}{\sum_{j=1}^n w_j(x_1, \dots, x_n; p_1, \dots, p_n)},$$

where $\{w_i\}_{i=1}^n$ is a set of positively valued weight functions that depend on

the lottery, $(x_1, \dots, x_n; p_1, \dots, p_n)$. A further restriction is obtained by imposing a desirable property called combination property: the utility of a lottery remains unchanged if we combine different probabilities of getting the same outcome. This implies that the w_i functions are of the form:

$$w_i(x_1, \dots, x_n; p_1, \dots, p_n) = p_i \tilde{w}_i(x_i; x_{-i}, p_{-i}),$$

where x_{-i} and p_{-i} denote the outcome vector and probability vector with the *i*th component deleted. With a bold sleight of hand, as yet unsubstantiated by any a priori reason, the \tilde{w}_i values are assumed to be obtained from a single function α evaluated at the *i*th outcome, x_i . This leads to a fairly tractable generalization of expected utility:

$$u(x_1, \dots, x_n; p_1, \dots, p_n) = \frac{\sum_{i=1}^n p_i \alpha(x_i)}{\sum_{j=1}^n p_j \alpha(x_j)} u(x_i).$$

Of course, a nice expression is just the first step. The next thing is to work back and forth in order to identify a minimal set of characteristic properties from which the α and u functions can be constructed. These properties, once found, would then necessarily be weaker than the corresponding ones for expected utility.

The first proof of such a representation theorem is in the context of generalizing the quasilinear mean of Hardy, Littlewood, and Polya given in Chapter 1. The corresponding result in terms of a preference ordering is given in Chapter 2. Although the basic result follows from a straightforward reinterpretation of those of Chapter 1, a self-contained treatment of Chapter 2 is presented so that readers who are used to the preference ordering approach can skip Chapter 1. Bob Weber provided an elegant geometrical interpretation of one of the axioms (Ratio Consistency) in an earlier version of the preference ordering theorem and demonstrated that it was redundant. This led to a

much weaker axiom in the current proof of the mean value representation theorem.

As is usual, the organization of the dissertation assumes the normal trappings of academic writing. Formal results, which are developed rigorously in Part I, precede interpretations in the context of decision theory in Part II, designed to build a *prima facie* case for the adoption of the more general expected utility hypothesis on both theoretical and empirical grounds. This division may cause the appearance of repetitiousness, but has the advantage of making Part II self-contained. Parts of the material in Part II are the result of joint work with Ken MacCrimmon, my research supervisor.

I would also like to acknowledge my debt to the other members on my guidance committee. The numerous instances when I went to Shelby Brumelle and Cindy Greenwood for help had been instrumental in enabling me to carry through the analysis in the proofs of the representation theorems, and also in the understanding of some basic mathematics. I have benefited from Daniel Kahneman's research with Amos Tversky on the psychology of judgement and decision-making under uncertainty and also from the seminars in his home. I have also benefited from discussions with Dave Donaldson on his work with Charles Blackorby on the measurement of inequality and poverty. I was exposed to research in the economics of uncertainty during Yoshitsugo Kanemoto's seminars. Thanks are also due to a non-member, John Butterworth, for suggesting the problem during the final examination of his information choice course.

To my grandmother

INTRODUCTION

Two well known representation theorems provide the starting point for this dissertation. The first, due to Hardy, Littlewood and Polya (1934), is an axiomatic characterization of a rather general class of mean values called the quasilinear mean:

$$M_{\phi}(F) = \phi^{-1}(\int_R \phi dF).$$

$M_{\phi}(F)$ denotes the quasilinear mean associated with a probability distribution F and characterized by a strictly monotone function ϕ . Hardy, Littlewood and Polya proved their characterization for simple probability distributions defined on a compact interval. Examples of quasilinear means include the widely used arithmetic mean (ϕ is linear), the geometric mean (ϕ is logarithmic), the harmonic mean (ϕ is of the form $\frac{1}{x}$) and the rth moment mean, also known as the general mean of order r (ϕ is of the form x^r).

The second representation theorem has its genesis in the St. Petersburg's paradox -- an individual is not willing to stake all that he possesses to take part in a lottery that pays 2^i dollars with $\frac{1}{2^i}$ chance; thus, demonstrating the limitation of using mathematical expectation of payoffs as a general rule for the ordering of risky prospects. This led Bernoulli to propose, in 1738, the expectation of a 'moral worth' function, u , of wealth as an alternative. In particular, he used the logarithmic function derived by assuming that an infinitesimal increase in the worth of wealth is proportional to an infinitesimal increase in wealth but inversely proportional to wealth itself. The monetary worth or certainty equivalent $M(F)$ correspond-

ing to a lottery F is then given by,

$$u(M(F)) = \int_{\mathbb{R}} u dF;$$

or alternatively,

$$M(F) = u^{-1}(\int_{\mathbb{R}} u dF).$$

Note that this expression is the same as that defining the quasilinear mean.

The first axiomatic treatment leading to the expectation of a function of payoffs as a rule for the ordering of lotteries is given by Ramsey (1926) in his "Foundation of Mathematics". von Neumann and Morgenstern (1947) provided an alternative axiomatization independently in their "Theory of Games and Economic Behavior" and initiated the use of the term "utility". They proved the existence of an order-preserving map on a mixture set (e.g. the space of probability distributions) subject to a minimal set of postulates such that the order-preserving map is the expectation of a utility function. Their result is now commonly referred to as the expected utility theorem.

The usefulness of quasilinear means needs no elaboration.¹ Expected utility theory, i.e., the application of the expected utility theorem to decision-making by interpreting the binary relation as a preference relation and the mixture set as a set of risky alternatives

¹ For a survey of the use of the r th moment mean in statistics, see Norris (1976). Blackorby & Donaldson (1978a) contains examples of the use of quasilinear mean in the measurement of income inequality. Weerahandi & Zidek (1979) provides an alternative characterization of the r th moment mean for probability distributions defined on the positive half-line. Ben-Tal (1977) showed that quasilinear means are ordinary arithmetic means defined on linear spaces with suitably chosen operations of addition and multiplication.

or equivalently, the interpretation of the quasilinear mean as a certainty equivalent, has attracted considerable attention since its inception. Alternative axiomatizations were given by Marschak (1950), Samuelson (1952), Herstein and Milnor (1953), Savage (1954), Anscombe and Aumann (1963), Pratt, Raiffa and Schlaifer (1964), Jensen (1967), DeGroot (1970), Fishburn (1970), Arrow (1971) and others. Savage (1954), Blackwell and Girshick (1961) and DeGroot (1970) applied expected utility theory to statistical decisions. In addition, it served as the foundation for Arrow (1971) and Marschak and Radner (1972) in their investigation of the economics of uncertainty, and for Howard (1964) and Keeney and Raiffa (1976) in their work on decision analysis.

Expected utility theory, though, has been less successful in describing and explaining actual choices (Edwards, 1961; Slovic, Fischhoff and Lichtenstein, 1977). Even von Neumann and Morgenstern realized at the outset that expected utility rules out complementarity among mutually exclusive consequences, a utility for gambling per se, and other behaviors that seem relatively common (von Neumann and Morgenstern, 1947, Appendix A, Sec. 3). Subsequently, various challenges, beginning with the Allais paradox (Allais, 1953), have called into question the empirical validity of a key property of expected utility theory, the strong independence principle (Marschak, 1950; Samuelson, 1952). The strong independence principle requires that ranking among lotteries remains unaltered when each lottery is composed with an identical lottery using the same probability. This is closely linked to the axiom of quasilinearity of Hardy, Littlewood and Polya (1934),

the "substitution of lotteries" of Pratt, Raiffa and Schlaifer (1964) and the "sure-thing" principle of Savage (1954).

Part I of this dissertation contains the statements and proofs of two representation theorems. The first theorem, proved in Chapter 1 generalizes the quasilinear mean of Hardy, Littlewood and Polya by weakening their axiom of quasilinearity. Given two distributions with the same means, quasilinearity requires that mixtures of these distributions with another distribution in the same proportions share the same mean, regardless of the distribution that they mixed with. We weaken the quasilinearity axiom by allowing the proportions that give rise to the same means to be different. This gives rise to a more general mean, $M_{\alpha\phi}$, that is specified by a continuous and strictly positive (negative) function, α , and a continuous and strictly monotone function, ϕ . The quasilinear mean results when the α function is constant. In addition, we show that the $M_{\alpha\phi}$ mean has the Intermediate Value Property, and provide necessary and sufficient conditions for consistency with the stochastic dominance (including higher degree ones) partial order. We also generalize a well known inequality among quasilinear means, by the observation that the $M_{\alpha\phi}$ mean of a distribution F can be written as the quasilinear mean of a distribution F^α where F^α is derived from F via α as the Radon-Nikodym derivative of F^α with respect to F .

As was noted earlier, the M_ϕ mean induces an ordering among distributions via the (expected utility) maximand,

$$\int_R \phi dF.$$

Correspondingly, the $M_{\alpha\phi}$ mean induces a more general ordering via the

maximand,

$$\int_R \alpha \phi dF / \int_R \alpha dF.$$

We prove, in Chapter 2, an alternative characterization of the above representation for simple probability measures on a more general outcome set where mean values may not be defined. In this case, axioms are stated directly in terms of properties of the underlying ordering. We retain several standard properties of expected utility, namely weak order, solvability and monotonicity but relax the substitution principle of Pratt, Raiffa and Schlaifer, which is essentially a restatement of quasilinearity in the context of an ordering.

The motivation for the research contained in Part I comes from paradoxes in the field of decision theory. Hence, the formulation of a new theory of choice (called alpha utility theory) that generalizes expected utility theory constitutes Part II (Chapters 3 and 4) of this dissertation. Chapter 3 gathers together the criticisms and empirical findings which contradict the implications of expected utility theory to pave the way for the development of alpha utility in the first two sections of Chapter 4. Sections 3 and 4 contain respectively, the derivation of the normative implications (e.g. stochastic dominance, global and local risk aversion) and descriptive implications (in particular, relating to the descriptive inadequacy of expected utility theory) of alpha utility theory. The chapter ends with a comparison in Section 5 of alpha utility with two alternative theories that have attracted significant interest.

PART I

TWO REPRESENTATION THEOREMS

GENERALIZING THE QUASILINEAR MEAN OF HARDY, LITTLEWOOD AND POLYA

1.1 INTRODUCTION

What is mean value? Conventional wisdom tells us that it represents, typifies or in some way measures the central tendency of a distribution. We are rescued from the ambiguity, as in elementary statistics texts, by examples. Some familiar notions of mean value include median, mode, arithmetic mean, geometric mean, harmonic mean and root-mean-square or more generally the r th root of the r th moment of a positive random variable (known also as the general mean of order r). Of these, the arithmetic mean is the most widely used. There are however situations for which the arithmetic mean may not be the most appropriate 'typical' value: notably, the discrepancy between per capita income and the 'typical' income for a society, the bulk of whose wealth is in the hands of a few.

In their pioneering work of 1934, Hardy, Littlewood and Polya showed how a general class of mean values, called the quasilinear mean, can be tailored to our needs subject to a necessary and sufficient set of axioms. This class includes as special cases all the examples of mean values mentioned above except for median and mode which do not satisfy their axioms. We generalize, in this chapter, the quasilinear mean via a weaker set of axioms stated in section 2. They are shown in section 3 to be necessary and sufficient for the representation of a class of mean values that generalizes the quasilinear mean. Finally, we derive some properties of our mean

value in section 4. Applications will be discussed in Part II and the conclusion of this dissertation.

1.2 AXIOMS OF MEAN VALUE²

Let D_J denote the space of probability distributions with all their mass concentrated in some interval J of the real line R (J need not be bounded). We consider a functional M whose domain is D_J . What properties should M possess in order to be a mean value? A natural candidate, motivated by the mean-value theorems of elementary calculus, is:

Property 1: Intermediate Value Property $M(F) \in \text{conv Supp}(F), \forall F \in D_J$.

The support of a distribution F , $\text{Supp}(F)$, consists of each point x such that every open set containing x has positive mass. $\text{Conv Supp}(F)$ is the smallest interval containing $\text{Supp}(F)$. The intermediate value property requires that the mean of a distribution be neither greater than the maximum attainable value nor less than the minimum attainable value. Axiom 1 below is a consequence of the intermediate value property.

Axiom 1: Consistency with Certainty $M(\delta_x) = x, \forall x \in J$,

The distribution δ_x refers to the step function at x which, in terms of probability, indicates obtaining x with probability 1.

Another property that seems reasonable is given by Axiom 2.

² In this section, the terms axiom and property are used interchangeably. Properties carry the "axiom" label if they appear in the final representation theorem as a characteristic property.

Axiom 2: Betweenness $\forall F, G \in D_J$, if $M(F) < M(G)$
 then $\forall \beta \in (0,1)$, $M(\beta F + (1-\beta)G) \in (M(F), M(G))$.

It is straightforward to check that Betweenness is equivalent to Property 2 stated below:

Property 2: Mixture-monotonicity $\forall F, G \in D_J$, if $M(F) < M(G)$, then
 $M(\beta F + (1-\beta)G) < M(\gamma F + (1-\gamma)G)$ if $1 \geq \beta \geq \gamma \geq 0$.

Lemma 1.1: Axiom 2 \Leftrightarrow Property 2.

Proof: Omitted.

A distribution G is said to stochastically dominate another distribution F in the first degree, denoted by $G \stackrel{1}{\geq} F$, if G is always not greater than F pointwise. If in addition, G is strictly less than F at some point, then G stochastically dominates F strictly in the first degree, denoted by $G \stackrel{1}{>} F$. Stochastic dominance of the first degree is an appealing partial order. Consistency of mean value with this partial order is stated as Property 3:

Property 3: Monotonicity $\forall F, G \in D_J$, $G \stackrel{1}{>} F \Rightarrow M(G) > M(F)$.

The next axiom deals with the effect on mean value of certain changes in the composition of the underlying distribution.

Axiom 3: Weak Substitution $\forall F, G \in D_J$, if $M(F) = M(G)$
 then $\forall \beta \in (0,1) \exists \gamma \in (0,1) \ni \forall H \in D_J$,
 $M(\beta F + (1-\beta)H) = M(\gamma G + (1-\gamma)H)$.

Hardy, Littlewood and Polya (1934) used a special case of Axiom 3 (They called it quasilinearity.) stated as Property 4 below for their quasilinear mean.

Property 4: Substitution (Quasilinearity) $\forall F, G, H \in D_J$, if $M(F)=M(G)$,
 then $\forall \beta \in (0,1)$
 $M(\beta F+(1-\beta)H)=M(\beta G+(1-\beta)H)$.

Starting with two distributions with the same mean value, quasilinearity or the substitution property requires that mixtures of these distributions with another distribution in the same proportions share the same mean regardless of the distribution that they mixed with. The weak substitution axiom allows mixture proportions that give rise to the same mean value to be different. The following property called Ratio Consistency is a consequence of Axioms 2 and 3 (In an earlier paper (Chew, 1979), the Ratio Consistency property was implicit in a stronger statement of Axiom 3. The weaker version is due to suggestions of Weber, Myerson, and Milgrom. For a discussion of an interesting geometrical interpretation of Ratio Consistency due to Weber, see Chapter 4, section 3.1):

Property 5: Ratio Consistency Suppose $F, G, H \in D_J$ and $\beta_1, \beta_2, \gamma_1, \gamma_2 \in (0,1)$
 $\ni M(F) = M(G) \neq M(H)$, and

$$M(\beta_i F+(1-\beta_i)H)=M(\gamma_i G+(1-\gamma_i)H) \quad \text{for } i=1,2.$$

$$\text{Then } \frac{\beta_1/1-\beta_1}{\gamma_1/1-\gamma_1} = \frac{\beta_2/1-\beta_2}{\gamma_2/1-\gamma_2}.$$

Lemma 1.2: Axioms 2 and 3 imply Property 5.

Proof: Suppose that $M(H) < M(F) = M(G)$ without loss of generality.

Axiom 3 $\Rightarrow \exists f: (0,1) \rightarrow (0,1) \ni \forall \beta \in (0,1),$

$$M(\beta F + (1-\beta)H) = M(f(\beta)G + (1-f(\beta))H) ;$$

Lemma 1.1 $\Rightarrow f$ is a strictly increasing function.

This together with Axioms 3 and 4 implies that f^{-1} exists.

Therefore, they are continuous functions, and hence differentiable a.e..

Define $\tau : (0,1) \rightarrow \mathbb{R}^+$ by $\tau(\beta) = \frac{f(\beta)/1-f(\beta)}{\beta/1-\beta}.$ (1.1)

Note that τ is continuous and differentiable a.e.. We show below that τ is a constant to complete the proof.

Consider $0 < \beta < \beta + \delta < 1$. It follows after substituting G for F using Axiom 3:

$$M((\beta+\delta)F + (1-(\beta+\delta))H) = M\left(\frac{(\beta+\delta)\tau(\beta+\delta)}{(\beta+\delta)\tau(\beta+\delta)+1-(\beta+\delta)} G + \frac{1-(\beta+\delta)}{(\beta+\delta)\tau(\beta+\delta)+1-(\beta+\delta)} H\right)$$
 (1.2)

$$\text{But } (\beta+\delta)F + (1-(\beta+\delta))H \equiv \beta F + (1-\beta) \left[\frac{\delta}{1-\beta} F + \frac{1-\beta-\delta}{1-\beta} H \right]$$

$$\text{Therefore L.H.S. of (1.2) } = M(\beta F + (1-\beta) \left[\frac{\delta}{1-\beta} F + \frac{1-\beta-\delta}{1-\beta} H \right])$$
 (1.3)

$$= M\left(\frac{\beta\tau(\beta)}{\beta\tau(\beta)+1-\beta} G + \frac{1-\beta}{\beta\tau(\beta)+1-\beta} \left[\frac{\delta}{1-\beta} F + \frac{1-\beta-\delta}{1-\beta} H \right]\right)$$
 (1.4)

after substituting G for F using Axiom 3.

Applying the same argument for the remaining F -component in expression (1.4), we obtain:

$$\begin{aligned} \text{R.H.S.} = M \left(\frac{\delta\tau(\delta/\beta\tau(\beta)+1-\beta)+\beta\tau(\beta)}{\delta\tau(\delta/\beta\tau(\beta)+1-\beta)+\beta\tau(\beta)+1-(\beta+\delta)} G^\circ \right. \\ \left. + \frac{1-(\beta+\delta)}{\delta\tau(\delta/\beta\tau(\beta)+1-\beta)+\beta\tau(\beta)+1-(\beta+\delta)} H \right). \end{aligned} \quad (1.5)$$

Comparing expressions (1.2) and expression (1.5), it follows that

$$(\beta+\delta)\tau(\beta+\delta) = \beta\tau(\beta) + \delta\tau(\delta/\beta\tau(\beta) + 1-\beta). \quad (1.6)$$

Suppose without loss of generality that τ is differentiable at β .

Then

$$\lim_{\delta \downarrow 0} \frac{(\beta+\delta)\tau(\beta+\delta) - \beta\tau(\beta)}{\delta} = \tau(\beta) + \beta\tau'(\beta) = \lim_{\delta \downarrow 0} \tau(\delta/\beta\tau(\beta)+1-\beta). \quad (1.7)$$

Therefore, the right hand limit of τ at 0 denoted by $\tau(0^+)$ exists.

Applying the same argument for other β 's for which τ is differentiable, we obtain

$$\frac{d}{d\beta}[\beta\tau(\beta)] = \tau(0^+) \quad \text{a.e.} \quad (1.8)$$

Therefore $\tau(\beta) = \tau(0^+) \quad \text{a.e.},$

and hence $\tau(\beta) \equiv \tau(0^+) \quad \text{by continuity.}$

Q.E.D.

Finally, we require our mean value to be a continuous functional in the sense of Axioms 4 and 5.

Axiom 4: Continuity If $\{F_n\}_{n=1}^\infty \subset D_J$ converges in distribution to

$F \in D_J$ and F has compact support, then

$$M(F) = \lim_{n \rightarrow \infty} M(F_n).$$

Convergence in distribution has the following characterization which is used sometimes as its definition.

F_n converges in distribution to $F \in D_J$ if and only if

$\int_J f dF_n$ converges to $\int_J f dF \quad \forall f \in C(J)$, where $C(J)$ is the space of all bounded continuous functions on J .

Note that when J is unbounded, the arithmetic mean of F_n does not necessarily converge to the arithmetic mean of F since the function x does not belong to $C(J)$. We impose the condition of compact support in Axiom 4 in order not to exclude the arithmetic mean from our class of mean values. The requirement of Continuity is useful because it tells us that the mean of a distribution may be approximated by the mean of a different distribution that is close to it. When J is a compact interval, Axiom 4 is equivalent to continuity of the mean value, M , with respect to the L^1 -norm. When J is unbounded, the following condition tells us how to estimate the mean value of a distribution F without compact support (if it exists) by its restriction to a compact interval, K , denoted by F_K .

Axiom 5: Extension Let $\{K_n\}_{n=1}^{\infty}$ be an increasing family of compact intervals

such that $\lim_{n \rightarrow \infty} K_n = J$, then $M(F) = \lim_{n \rightarrow \infty} M(F_{K_n}), \forall F \in D_J$.

The mean value for a distribution F without compact support is given by the limit of the mean values of the sequence of truncated distributions, $\{F_{K_n}\}_{n=1}^{\infty}$, for any increasing family, $\{K_n\}_{n=0}^{\infty}$, whose limit is J . Since the sequence of mean values, $\{M(F_{K_n})\}_{n=0}^{\infty}$ does not always converge, the mean value for a distribution without compact support need not exist. A good example is the arithmetic mean.

1.3 REPRESENTATION THEOREM

We begin with a statement of the quasilinear mean representation theorem. $D^0[A,B]$ denotes the restriction of $D[A,B]$ to simple distributions with a finite number of discontinuities.

Theorem 1.1: (Hardy, Littlewood & Polya)

Suppose $\exists M : D^0[A,B] \rightarrow R$.

Then M satisfies Axiom 1, Property 3 and Property 4,

if and only if $\exists \phi : [A,B] \rightarrow R$, continuous, strictly monotone such that $M(F) = \phi^{-1} \left(\int_A^B \phi \, dF \right), \quad \forall F \in D^0[A,B].$ (1.9)

Moreover, if $\exists \phi^* : [A,B] \rightarrow R$

such that $M(F) = \phi^{*-1} \left(\int_A^B \phi^* \, dF \right), \quad \forall F \in D^0[A,B]$

then $\forall x \in [A,B], \phi^*(x) = a\phi(x) + b$, for some a, b with $a \neq 0$.

(1.10)

Proof: (Omitted since it is a special case of Theorem 1.2).

In other words, the most general *certainty consistent, monotone and quasilinear* functional of F is that defined by (1.9). We shall call it *quasilinear mean*. Since the quasilinear mean $M(F)$ is completely specified by a continuous, strictly increasing function ϕ , up to an affine transformation, (1.10), it is convenient to write it as $M_\phi(F)$. The following theorem generalizes the theorem of Hardy, Littlewood and Polya and extends their analysis from $D^0[A,B]$ to $D[A,B]$. A further extension to D_J for the arbitrary J follows later.

Theorem 1.2: Suppose $\exists M: D[A,B] \rightarrow R$.

Then M satisfies Axiom 1, Axiom 2, Axiom 3 and Axiom 4

if and only if $\exists \phi: [A,B] \rightarrow R$, continuous, strictly monotone,

and $\alpha: [A,B] \rightarrow R^+$, continuous, strictly positive,

such that $M(F) = \phi^{-1}(\int_A^B \alpha \phi dF / \int_A^B \alpha dF)$, $\forall F \in D[A,B]$. (1.11)

Moreover, if $\exists \phi^*: [A,B] \rightarrow R$, and $\alpha^*: [A,B] \rightarrow R^+$,

such that $M(F) = \phi^{*-1}(\int_A^B \alpha^* \phi^* dF / \int_A^B \alpha^* dF)$, $\forall F \in D[A,B]$

then $\forall x \in [A,B]$, $\phi^*(x) = a \frac{k(\phi(x)-\phi(A))}{k(\phi(x)-\phi(A)) + (\phi(B)-\phi(x))} + b$, (1.12)

and $\alpha^*(x) = c\alpha(x)\{k(\phi(x)-\phi(A)) + (\phi(B)-\phi(x))\}$, (1.13)

for some a,b,c,k with $a,c \neq 0$, $k > 0$.

Proof: (Necessity)

Axiom 1 follows immediately.

Axiom 2 follows from the observation that,

$$\phi(M(\beta F + (1-\beta)G)) = \frac{\beta(\int_A^B \alpha dF) \phi(M(F)) + (1-\beta)(\int_A^B \alpha dG) \phi(M(G))}{\beta(\int_A^B \alpha dF) + (1-\beta)(\int_A^B \alpha dG)}$$

increases strictly in β when $M(F) > M(G)$.

Consider $F,G,H \in D[A,B]$, Suppose $M(F) = M(G)$.

$$\text{Then, } \phi(M(\beta F + (1-\beta)H)) = \frac{\beta(\int_A^B \alpha dF) \phi(M(F)) + (1-\beta)(\int_A^B \alpha dH) \phi(M(H))}{\beta(\int_A^B \alpha dF) + (1-\beta)(\int_A^B \alpha dH)}$$

$$= \phi(M(\gamma G + (1-\gamma)H))$$

$$\text{with } \frac{\gamma/1-\gamma}{\beta/1-\beta} = (\int_A^B \alpha dF) / (\int_A^B \alpha dG), \forall \beta \in [A,B].$$

Hence, Axiom 3.

Let $\{F_n\}_{n=1}^{\infty}$ converge to F . Since α, ϕ, ϕ^{-1} are continuous on a compact interval $[A, B]$,

$$\begin{aligned} \int_A^B \alpha \, dF_n &\rightarrow \int_A^B \alpha \, dF \quad \text{and} \quad \int_A^B \alpha \phi \, dF_n \rightarrow \int_A^B \alpha \phi \, dF . \\ \Rightarrow \int_A^B \alpha \phi \, dF_n / \int_A^B \alpha \, dF_n &\rightarrow \int_A^B \alpha \phi \, dF / \int_A^B \alpha \, dF . \\ \Rightarrow M(F_n) = \phi^{-1}(\int_A^B \alpha \phi \, dF_n / \int_A^B \alpha \, dF_n) &\rightarrow \phi^{-1}(\int_A^B \alpha \phi \, dF / \int_A^B \alpha \, dF) = M(F) . \end{aligned}$$

Axiom 4 follows.

(Sufficiency)

Define $\psi : [0, 1] \rightarrow [A, B]$ as follows.

$$\psi(p) = M(S_p), \quad \forall p \in [0, 1], \quad \text{where } S_p \equiv p\delta_B + (1-p)\delta_A. \quad (1.14)$$

Axiom 1 $\Rightarrow \psi(0) = A$ and $\psi(1) = B$.

Axiom 2 $\Rightarrow \psi$ is strictly increasing.

Let $\{p_n\}_{n=1}^{\infty}$ converge to p . Then S_{p_n} converges in distribution to S_p .

$$\text{Axiom 4} \Rightarrow \psi(p) = M(S_p) = \lim_{n \rightarrow \infty} M(S_{p_n}) = \lim_{n \rightarrow \infty} \psi(p_n).$$

It follows that ψ is continuous and strictly increasing and there-

fore has an inverse $\tilde{\phi} : [A, B] \rightarrow [0, 1]$ which is continuous and

strictly increasing. If $x = \psi(p)$, then $p = \tilde{\phi}(x)$

$$\text{and} \quad M(\delta_x) = x = \psi(p) = M(S_p) = M(S_{\tilde{\phi}(x)}). \quad (1.15)$$

Lemma 1.2 implies the existence of a strictly positive constant τ_x

that depends on x such that, $\forall H \in D[A, B], \forall \beta \in [0, 1]$,

$$M(\beta\delta_x + (1-\beta)H) = M\left(\frac{\beta\tau_x}{\beta\tau_x + (1-\beta)} S_{\tilde{\phi}(x)} + \frac{1-\beta}{\beta\tau_x + (1-\beta)} H\right). \quad (1.16)$$

Construct $\tilde{\alpha}: (A,B) \rightarrow (0,\infty)$ by assigning $\tilde{\alpha}(x) = \tau_x$, $\forall x \in (A,B)$.

The following argument establishes the continuity of $\tilde{\alpha}$ on (A,B) and then extends its domain to include the end-points.

$$\text{Consider } g(x) = M(\frac{1}{2}\delta_x + \frac{1}{2}\delta_A) = M(\frac{\tilde{\alpha}(x)}{\tilde{\alpha}(x)+1} S_{\tilde{\phi}(x)} + \frac{1}{\tilde{\alpha}(x)+1} \delta_A)$$

$$= M(S_{\{\tilde{\alpha}(x)\tilde{\phi}(x)/\tilde{\alpha}(x)+1\}}) = \psi(\tilde{\alpha}(x)\tilde{\phi}(x)/\tilde{\alpha}(x)+1).$$

Let $\{x_n\}_{n=1}^{\infty}$ converge to $x \in (A,B)$. Then $\frac{1}{2}\delta_{x_n} + \frac{1}{2}\delta_A$ converges in distribution to $\frac{1}{2}\delta_x + \frac{1}{2}\delta_A$. Axiom 4 implies that,

$$g(x) = M(\frac{1}{2}\delta_x + \frac{1}{2}\delta_A) = \lim_{n \rightarrow \infty} M(\frac{1}{2}\delta_{x_n} + \frac{1}{2}\delta_A) = \lim_{n \rightarrow \infty} g(x_n).$$

Therefore, g is continuous in (A,B) .

It follows that $\tilde{\alpha}$ is continuous in (A,B) .

Let $\{x_n\}_{n=1}^{\infty}$ converge to B from below, then,

$$\begin{aligned} \psi(\frac{1}{2}) &= M(S_{\frac{1}{2}}) = \lim_{n \rightarrow \infty} M(\frac{1}{2}\delta_{x_n} + \frac{1}{2}\delta_A) \\ &= \lim_{n \rightarrow \infty} g(x_n) = \lim_{n \rightarrow \infty} \psi(\tilde{\phi}(x_n)/(1+1/\tilde{\alpha}(x_n))). \end{aligned}$$

$$\Rightarrow \lim_{n \rightarrow \infty} \tilde{\alpha}(x_n) = 1, \text{ since } \tilde{\phi}(B) = 1.$$

Similarly, we can show that $\lim_{n \rightarrow \infty} \tilde{\alpha}(x_n) = 1$ as x_n converges to A from above. We extend $\tilde{\alpha}$ to $[A,B]$ continuously by assigning 1 to $\tilde{\alpha}$ at the end-points. Now, we are ready to show that, $\forall F \in D^0[A,B]$, the functions $\tilde{\alpha}$ and $\tilde{\phi}$ satisfy condition (1.11).

Let $\{x_i\}_{i=1}^n$ be the support of a distribution F in $D^0[A,B]$, and

represent F in the form, $F \equiv \sum_{i=1}^n \theta_i \delta_{x_i}$, where $\theta_i = F(x_i) - F(x_i^-)$.

$$\begin{aligned}
M(F) &= M\left(\sum_{i=1}^n \theta_i \delta_{x_i}\right) \\
&= M\left(\frac{\theta_1 \tilde{\alpha}(x_1)}{\theta_1 \tilde{\alpha}(x_1) + \sum \theta_j} S_{\tilde{\phi}(x_1)} + \sum_{i=2}^n \frac{\theta_i}{\theta_1 \tilde{\alpha}(x_1) + \sum \theta_j} \delta_{x_i}\right).
\end{aligned}$$

after substituting $S_{\tilde{\phi}(x_1)}$ for δ_{x_1} using expression (1.16).

Repeating (n-1) times on the remaining δ_{x_i} , $i=2, \dots, n$, yields,

$$\begin{aligned}
M(F) &= M\left(\sum_{i=1}^n \left\{\frac{\theta_i \tilde{\alpha}(x_i)}{\sum \theta_j \tilde{\alpha}(x_j)}\right\} S_{\tilde{\phi}(x_i)}\right) = M(S_{\{\sum \theta_i \tilde{\alpha}(x_i) \tilde{\phi}(x_i) / \sum \theta_j \tilde{\alpha}(x_j)\}}) \\
&= \psi(\sum \theta_i \tilde{\alpha}(x_i) \tilde{\phi}(x_i) / \sum \theta_j \tilde{\alpha}(x_j)) = \tilde{\phi}^{-1} \left(\int_A^B \tilde{\alpha} \tilde{\phi} dF / \int_A^B \tilde{\alpha} dF \right).
\end{aligned}$$

Finally, we extend our construction to $F \in D[A, B]$.

Suppose $F \in D[A, B] - D^0[A, B]$. Construct the following sequence

$\{F_n\}_{n=1}^\infty$ in $D^0[A, B]$.

$$F_n \equiv F(A) \delta_A + \sum_{i=1}^{2^n} \left\{ F\left(A + \frac{i(B-A)}{2^n}\right) - F\left(A + \frac{(i-1)(B-A)}{2^n}\right) \right\} \delta_{A+i(B-A)/2^n}$$

By construction, $F_n(x) \rightarrow F(x)$, $\forall x \in \{A + \frac{i(B-A)}{2^n} : i, n \in \mathbb{I}^+, i \leq 2^n\}$,

which is dense in $[A, B]$. Therefore $\{F_n\}_{n=1}^\infty$ converges to F .

$$\text{Axiom 4} \Rightarrow M(F) = \lim_{n \rightarrow \infty} M(F_n)$$

$$\begin{aligned}
&= \lim_{n \rightarrow \infty} \tilde{\phi}^{-1} \left(\int_A^B \tilde{\alpha} \tilde{\phi} dF_n / \int_A^B \tilde{\alpha} dF_n \right) \\
&= \tilde{\phi}^{-1} \left(\int_A^B \tilde{\alpha} \tilde{\phi} dF / \int_A^B \tilde{\alpha} dF \right),
\end{aligned}$$

since $\tilde{\phi}, \tilde{\phi}^{-1}$ and $\tilde{\alpha}$ are continuous on $[A, B]$.

(Uniqueness)

Suppose $\exists \alpha : [A, B] \rightarrow \mathbb{R}^+$ and $\phi : [A, B] \rightarrow \mathbb{R}$, that satisfy condition (1.11).

Then $x = M(\delta_x) = M(S_{\tilde{\phi}(x)})$

$$= \phi^{-1} \left(\frac{\alpha(B)\phi(B)\tilde{\phi}(x) + \alpha(A)\phi(A)(1-\tilde{\phi}(x))}{\alpha(B)\tilde{\phi}(x) + \alpha(A)(1-\tilde{\phi}(x))} \right). \quad (1.17)$$

and $\forall \beta \in (0, 1)$,

$$\begin{aligned} & \phi^{-1} \left(\frac{\beta\alpha(x)\phi(x) + (1-\beta)\alpha(A)\phi(A)}{\beta\alpha(x) + (1-\beta)\alpha(A)} \right) \\ &= M(\beta\delta_x + (1-\beta)\delta_A) \\ &= M\left(\frac{\beta\tilde{\alpha}(x)}{\beta\tilde{\alpha}(x) + (1-\beta)} S_{\tilde{\phi}(x)} + \frac{(1-\beta)}{\beta\tilde{\alpha}(x) + (1-\beta)} \delta_A\right) \\ &= \phi^{-1} \left(\frac{\beta\tilde{\alpha}(x)\tilde{\phi}(x)\alpha(B)\phi(B) + (\beta\tilde{\alpha}(x)(1-\tilde{\phi}(x)) + (1-\beta))\alpha(A)\phi(A)}{\beta\tilde{\alpha}(x)\tilde{\phi}(x)\alpha(B) + (\beta\tilde{\alpha}(x)(1-\tilde{\phi}(x)) + (1-\beta))\alpha(A)} \right), \end{aligned} \quad (1.18)$$

after applying α and ϕ to the equalities (1.15) and (1.16).

Let $a = \phi(B) - \phi(A)$,

$b = \phi(A)$,

$c = \alpha(A)$,

and $k = \alpha(B)/\alpha(A)$.

It is straightforward to check that, $\forall x \in [A, B]$,

$$\phi(x) = a\{k\tilde{\phi}(x)/(k\tilde{\phi}(x) + (1-\tilde{\phi}(x)))\} + b,$$

$$\text{and} \quad \alpha(x) = c\tilde{\alpha}(x)\{k\tilde{\phi}(x) + (1-\tilde{\phi}(x))\}.$$

Suppose α^* and ϕ^* are another pair of functions that satisfy condition (1.11).

Then $\phi^*(x) = a^*\{k^*\tilde{\phi}(x)/(k^*\tilde{\phi}(x) + (1-\tilde{\phi}(x)))\} + b^*$,

$$\text{and } \alpha^*(x) = c^* \tilde{\alpha}(x) \{k^* \tilde{\phi}(x) + (1 - \tilde{\phi}(x))\},$$

$$\text{with } a^* = \phi^*(B) - \phi^*(A),$$

$$b^* = \phi^*(A),$$

$$c^* = \alpha^*(A),$$

$$\text{and } k^* = \alpha^*(B)/\alpha^*(A).$$

Finally, we check that, $\forall x \in [A, B]$,

$$\phi^*(x) = a' \{k' (\phi(x) - \phi(A)) / (k' (\phi(x) - \phi(A)) + (\phi(B) - \phi(x)))\} + b',$$

$$\alpha^*(x) = c' \alpha(x) \{k' (\phi(x) - \phi(A)) + (\phi(B) - \phi(x))\},$$

$$\text{for } a' = (\phi^*(B) - \phi^*(A)) / a = a^* / a,$$

$$b' = \phi^*(A) = b^*,$$

$$c' = \alpha^*(A) / \{\alpha(A) (\phi(B) - \phi(A))\} = c^* / ca,$$

$$\text{and } k' = \{\alpha^*(B) \alpha(A) / \alpha^*(A) \alpha(B)\} = k^* / k.$$

Q.E.D.

Our generalization of the quasilinear mean, defined by (1.11), is completely specified by a pair of functions (α, ϕ) . According to Theorem 1.2, this is the most general mean for distributions defined on a compact interval that satisfies *Consistency with Certainty*, *Betweenness*, *Weak Substitution* and *Continuity*. In keeping with precedent, we denote our generalized mean by $M_{\alpha\phi}$. The pair of functions (α, ϕ) denotes a particular member of the class, $\{\alpha, \phi\}_J$, of functions that yield the same mean on D_J . When J is a compact interval, such as $[A, B]$ in Theorem 1.2, we can form the following subclass of $\{\alpha, \phi\}_J$, called *k-ratio subclass*, for $k > 0$.

$$\{\alpha, \phi\}_{[A, B]}^k \equiv \{(\alpha, \phi) \in \{\alpha, \phi\}_{[A, B]} : \alpha(B)/\alpha(A) = k\}.$$

We denote by (α^k, ϕ^k) a generic element and $(\tilde{\alpha}^k, \tilde{\phi}^k)_{[A,B]}$ the *canonical* element of $\{\alpha, \phi\}_{[A,B]}^k$ that satisfies: $\tilde{\phi}^k(A)=0$, $\tilde{\phi}^k(B)=1$, $\tilde{\alpha}^k(A)=1$ and $\tilde{\alpha}^k(B)=k$. It can be shown, using expressions (1.12) and (1.13), that the elements of a k-ratio subclass are related to each other via an affine transformation for the ϕ^k component and a scalar transformation for the α^k component. The class $\{\alpha, \phi\}_{[A,B]}$ can be obtained from its k-ratio subclasses by taking their union over all positive k's.

The following corollaries of Theorem 1.2 are needed to extend our results to include noncompact intervals. The restriction of $\{\alpha, \phi\}_{[A_1, B_1]}$ to the interval $[A_0, B_0]$ is denoted by $\{\alpha, \phi\}_{[A_1, B_1]}|_{[A_0, B_0]}$.

Corollary 1.1: Let $A_1 < A_0 < B_0 < B_1$.

$$\text{Then } \{\alpha, \phi\}_{[A_1, B_1]}|_{[A_0, B_0]} = \bigcup_{k \in (h_{01}, \bar{h}_{01})} \{\alpha, \phi\}_{[A_0, B_0]}^k,$$

$$\text{where } \bar{h}_{01} = \frac{\tilde{\alpha}^1(B_0)}{\tilde{\alpha}^1(A_0)} \frac{\tilde{\phi}^1(B_0)}{\tilde{\phi}^1(A_0)}, \quad (1.19)$$

$$\text{and } h_{01} = \frac{\tilde{\alpha}^1(B_0)}{\tilde{\alpha}^1(A_0)} \frac{(1 - \tilde{\phi}^1(B_0))}{(1 - \tilde{\phi}^1(A_0))}. \quad (1.20)$$

Proof: Denote by $(\alpha_1^{k_1}, \phi_1^{k_1})$ an element of $\{\alpha, \phi\}_{[A_1, B_1]}^{k_1}$.

Observe that $(\alpha_1^{k_1}, \phi_1^{k_1})|_{[A_0, B_0]} \in \{\alpha, \phi\}_{[A_0, B_0]}^{k_0}$,

$$\text{where } k_0 = \frac{\alpha_1^{k_1}(B_0)}{\alpha_1^{k_1}(A_0)} = \frac{\tilde{\alpha}_1^{k_1}(B_0)(k_1 \tilde{\phi}_1^{k_1}(B_0) + (1 - \tilde{\phi}_1^{k_1}(B_0)))}{\tilde{\alpha}_1^{k_1}(A_0)(k_1 \tilde{\phi}_1^{k_1}(A_0) + (1 - \tilde{\phi}_1^{k_1}(A_0)))}, \quad (1.21)$$

which is a continuous, strictly increasing and onto function of k_1

with domain $(0, \infty)$ and range (h_{01}, \bar{h}_{01}) .

Q.E.D.

Corollary 1.2: Let $A_2 < A_1 < A_0 < B_0 < B_1 < B_2$.

Denote by $(\tilde{\alpha}_i^{-1}, \tilde{\phi}_i^{-1})$ the canonical element of the unitary subclass.

for the interval $[A_i, B_i]$, $i=1,2$.

Then $h_{02} < h_{01}$ and $\bar{h}_{02} > \bar{h}_{01}$, (1.22)

where $\bar{h}_{0i} = \frac{\tilde{\alpha}_i^{-1}(B_0) \tilde{\phi}_i^{-1}(B_0)}{\tilde{\alpha}_i^{-1}(A_0) \tilde{\phi}_i^{-1}(A_0)}$,

and $h_{0i} = \frac{\tilde{\alpha}_i^{-1}(B_0)(1 - \tilde{\phi}_i^{-1}(B_0))}{\tilde{\alpha}_i^{-1}(A_0)(1 - \tilde{\phi}_i^{-1}(A_0))}$, for $i = 1, 2$.

Proof: Construct the functions $\xi_{j,i} : (0, \infty) \rightarrow (h_{ij}, \bar{h}_{ij})$, for $i < j$,

via $k_i = \xi_{j,i}(k_j)$

$$= \frac{\tilde{\alpha}_j^{-1}(B_i)(k_j \tilde{\phi}_j^{-1}(B_i) + (1 - \tilde{\phi}_j^{-1}(B_i)))}{\tilde{\alpha}_j^{-1}(A_i)(k_j \tilde{\phi}_j^{-1}(A_i) + (1 - \tilde{\phi}_j^{-1}(A_i)))}. \quad (1.23)$$

Note that, by construction,

$$\{\alpha, \phi\}_{[A_j, B_j]}^{k_j} \Big|_{[A_i, B_i]} = \{\alpha, \phi\}_{[A_i, B_i]}^{\xi_{j,i}(k_j)}$$

Note also that $\xi_{j,i}$ is continuous, strictly increasing, and onto from $(0, \infty)$ to (h_{ij}, \bar{h}_{ij}) .

Suppose $\bar{h}_{02} \geq \bar{h}_{01}$.

Pick $k_1 \in [\bar{h}_{12}, \infty)$. Then $\xi_{10}(k_1) < \bar{h}_{01} \leq \bar{h}_{02}$.

$\Rightarrow \exists k_2 \in (0, \infty)$ such that $\xi_{20}(k_2) = \xi_{10}(k_1)$.

But $\{\alpha, \phi\}_{[A_2, B_2]}^{k_2} \Big| [A_0, B_0]$

$$= \{\alpha, \phi\}_{[A_2, B_2]}^{k_2} \Big| [A_1, B_1] \Big| [A_0, B_0]$$

$$= \{\alpha, \phi\}_{[A_1, B_1]}^{\xi_{21}(k_2)} \Big| [A_0, B_0]$$

$$\Rightarrow k_1 = \xi_{21}(k_2) \in (\underline{h}_{12}, \bar{h}_{12}) \quad \neq$$

A similar argument establishes $\underline{h}_{01} < \underline{h}_{02}$.

Q.E.D.

According to Corollary 1.1, we have to restrict the k_0 -ratio corresponding to the $[A_0, B_0]$ interval to within a range of values if we want $(\alpha_0^{k_0}, \phi_0^{k_0})$ to agree with $(\alpha_1^{k_1}, \phi_1^{k_1})$ restricted to $[A_0, B_0]$ for a larger interval $[A_1, B_1]$. Corollary 1.2 tells us that the range of permissible k_0 -ratio's gets squeezed as we go from $[A_1, B_1]$ to a larger interval $[A_2, B_2]$. Now we extend Theorem 1.2 to the case of arbitrary interval J .

Theorem 1.3: Suppose $\exists M : D_J \rightarrow \mathbb{R}$.

Then M satisfies Axiom 1, Axiom 2, Axiom 3, Axiom 4 and Axiom 5

if and only if $\exists \phi : J \rightarrow \mathbb{R}$, continuous, strictly monotone,

and $\alpha : J \rightarrow \mathbb{R}^+$, continuous, strictly positive,

such that ${}^3 M(F) = \phi^{-1}(\int_J \alpha \phi dF / \int_J \alpha dF)$, $\forall F \in D_J$. (1.24)

Moreover, if (α^*, ϕ^*) is another pair of functions that satisfies

³ The ratio $\int_J \alpha \phi dF / \int_J \alpha dF$ for F without compact support is defined by expression (1.25).

condition (1.24), then \forall interval $[A, B] \subset J$, $\exists a, b, c, k$ with $a, c \neq 0$ and $k > 0 \ni \forall x \in [A, B]$,

$$\phi^*(x) = a \frac{k(\phi(x) - \phi(A))}{k(\phi(x) - \phi(A)) + (\phi(B) - \phi(x))} + b,$$

$$\alpha^*(x) = c\alpha(x)\{k(\phi(x) - \phi(A)) + (\phi(B) - \phi(x))\}.$$

Proof: (Necessity)

Verification of Axiom 1, Axiom 2, Axiom 3, and Axiom 4 is the same as in proof of Theorem 1.2. Axiom 5 follows trivially from the definition (expression (1.25)) of $M(F)$ for F without compact support.

(Sufficiency)

If J is compact, then we are done. Otherwise, let

$\{K_n\}_{n=0}^{\infty} = \{[A_n, B_n]\}_{n=0}^{\infty}$ be a sequence of intervals such that

$\{A_n\}_{n=0}^{\infty}$ ($\{B_n\}_{n=0}^{\infty}$) is a strictly decreasing (increasing) sequence, and $\lim_{n \rightarrow \infty} K_n = J$.

Corollary 1.1 says that,

$$\{\alpha_i, \phi_i\}_{[A_i, B_i]} \Big|_{[A_0, B_0]} = \bigcup_{i=0}^{\infty} k_0^{\varepsilon}(h_{-0i}, \bar{h}_{0i})^{\{\alpha_0, \phi_0\}} [A_0, B_0]^{k_0},$$

$$\text{where } (h_{-0i}, \bar{h}_{0i}) = \left(\frac{\tilde{\alpha}_i^{-1}(B_0) (1 - \tilde{\phi}_i^{-1}(B_0))}{\tilde{\alpha}_i^{-1}(A_0) (1 - \tilde{\phi}_i^{-1}(A_0))}, \frac{\tilde{\alpha}_i^{-1}(B_0) \tilde{\phi}_i^{-1}(B_0)}{\tilde{\alpha}_i^{-1}(A_0) \tilde{\phi}_i^{-1}(A_0)} \right)$$

Corollary 1.2 says that,

$\{h_{-0i}\}_{i=1}^{\infty}$ is strictly increasing and $\{\bar{h}_{0i}\}_{i=1}^{\infty}$ is strictly decreasing.

Let $C_i = [(h_{-0i} + h_{-0\ i+1})/2, (\bar{h}_{0i} + \bar{h}_{0\ i+1})/2]$,

and $D_i = (h_{-0\ i+1}, \bar{h}_{0\ i+1})$.

Then $D_i \subsetneq C_i \subsetneq (h_{-0i}, \bar{h}_{0i})$ for $i = 1, 2, 3, \dots$.

Observe that (h_{-0i}, \bar{h}_{0i}) , C_i, D_i are strictly decreasing sequences of sets by inclusion. Since C_i is compact for each i ,

therefore, $\lim_{i \rightarrow \infty} C_i = C_\infty \neq \emptyset$.

Since $D_i \subsetneq C_i \subsetneq (h_{-0i}, \bar{h}_{0i}) \forall i$, it follows that,

$$\lim_{i \rightarrow \infty} D_i \subset C_\infty \subset \lim_{i \rightarrow \infty} (h_{-0i}, \bar{h}_{0i}).$$

But $\lim_{i \rightarrow \infty} D_i = \lim_{i \rightarrow \infty} (h_{-0i}, \bar{h}_{0i})$.

Hence $\lim_{i \rightarrow \infty} (h_{-0i}, \bar{h}_{0i}) = C_\infty$.

To construct (α, ϕ) defined on J that satisfies condition (1.24),

pick $k_0 \in C_\infty$.

$$\begin{aligned} \text{Define } (\alpha(x), \phi(x)) &= (\alpha_0^{k_0}(x), \phi_0^{k_0}(x)) \quad \text{for } x \in [A_0, B_0], \\ &= (\alpha_1^{k_1}(x), \phi_1^{k_1}(x)) \quad \text{for } x \in [A_1, B_1] - [A_0, B_0], \\ &\quad \vdots \\ &= (\alpha_i^{k_i}(x), \phi_i^{k_i}(x)) \quad \text{for } x \in [A_i, B_i] - [A_{i-1}, B_{i-1}], \\ &\quad \vdots \end{aligned}$$

such that $\xi_{i0}(k_i) = k_0 = \tilde{\alpha}_0^{k_0}(B_0)$, (see expression (1.23)),

$$\alpha_{ii}^{k_i}(A_0) = 1 = \tilde{\alpha}_0^{k_0}(A_0),$$

$$\phi_{ii}^{k_i}(A_0) = 0 = \tilde{\phi}_0^{k_0}(A_0),$$

$$\phi_i^{k_i}(B_0) = 1 = \phi_0^{k_0}(B_0).$$

Observe that $(\alpha_{i+1}^{k_{i+1}}, \phi_{i+1}^{k_{i+1}}) \Big|_{[A_i, B_i]} \equiv (\alpha_i^{k_i}, \phi_i^{k_i})$ since they agree at A_0 and at B_0 .

Given any distribution F with compact support, pick $[A_i, B_i]$ such that $\text{Supp}(F) \subset [A_i, B_i]$.

$$\begin{aligned} \text{Then } M(F) &= \phi_i^{k_i}{}^{-1} \left(\int_{A_i}^{B_i} \alpha_i^{k_i} \phi_i^{k_i} dF / \int_{A_i}^{B_i} \alpha_i^{k_i} dF \right), \\ &= \phi^{-1} \left(\int_J \alpha \phi dF / \int_J \alpha dF \right). \end{aligned}$$

For any distribution $F \in D_J$, if $\text{Supp}(F)$ is not compact, then we obtain $M(F)$, if it exists, from Axiom 5 as follows.

Let $\{K_n\}_{n=0}^{\infty}$ be an increasing sequence of compact intervals whose limit is equal to J . Denote by F_{K_n} the restriction of F to K_n .

$$\begin{aligned} \text{Then Axiom 5 } \Rightarrow M(F) &= \lim_{n \rightarrow \infty} M(F_{K_n}) \\ &= \lim_{n \rightarrow \infty} \phi^{-1} \left(\int_J \alpha \phi dF_{K_n} / \int_J \alpha dF_{K_n} \right). \end{aligned} \quad (1.25)$$

When the limit (1.25) exists and does not depend on the choice of the sequence, $\{K_n\}_{n=0}^{\infty}$, it is denoted by $\phi^{-1}(\int_J \alpha \phi dF / \int_J \alpha dF)$. However, the above limit does not always exist. An example is given by the arithmetic mean for a cauchy distribution.

(Uniqueness)

This follows directly from applying Theorem 1.2 to arbitrary intervals $[A, B]$ in J .

Q.E.D.

We have characterized the class of mean values for distributions on the real line having the properties of *Consistency with Certainty*, *Betweenness*, *Weak Substitution*, *Continuity* and *Extension* with a pair of functions (α, ϕ) . For distributions without compact supports, their corresponding $M_{\alpha\phi}$ means do not necessarily exist (see (1.25)). A necessary and sufficient condition that ensures existence is given below.

Corollary 1.3: $M_{\alpha\phi}(F)$ exists $\forall F \in D_J$ if and only if either ϕ is bounded or $\alpha \cdot \phi$ is bounded.

Proof: The sufficiency part of the proof is straightforward.

To prove necessity, suppose for the pair (α, ϕ) , neither ϕ nor $\alpha \cdot \phi$ is bounded. We may assume, without losing generality, that ϕ is not bounded from above. There are two cases. As x tends to $+\infty$, either i) $\alpha(x)$ is bounded from above, or ii) $\alpha(x)$ tends to $+\infty$.

Case i): Consider a sequence $\{x_i\}_{i=1}^{\infty} \ni \alpha(x_i)\phi(x_i) = 2^i$.

Then $M_{\alpha\phi}(\sum_{i=1}^{\infty} \frac{1}{2^i} \delta_{x_i}) = \lim_{m \rightarrow \infty} \frac{m}{\sum_{i=1}^m \alpha(x_i)/2^i}$ does not converge.

Case ii): Consider a sequence $\{x_i\}_{i=1}^{\infty} \ni \phi(x_i) = 2^i$.

Then $M_{\alpha\phi}(\sum_{i=1}^{\infty} \frac{1}{2^i} \delta_{x_i}) = \lim_{m \rightarrow \infty} \frac{\sum_{i=1}^m \alpha(x_i)}{\sum_{i=1}^m \alpha(x_i)/2^i}$ does not converge.

A similar argument establishes the result for the case when ϕ is unbounded from below.

Q.E.D.

The above corollary is useful in Chapter 4 when we interpret mean value as the certainty equivalent of a lottery and insist that a certainty equivalent should always be finite.

1.4 PROPERTIES OF THE $M_{\alpha\phi}$ MEAN

Of possible properties for mean value, the Intermediate Value property (Property 1) enjoys a rather special status, somewhat like a defining property. After all, even measures such as median and mode, which are rejects of the quasilinear mean, exhibit this property. The conclusion that $M_{\alpha\phi}$ has the intermediate value property follows from the observation that,

$$M_{\alpha\phi}(F) = c \Rightarrow \frac{\int \alpha(x) (\phi(x) - \phi(c)) dF(x)}{\int \alpha(x) dF(x)} = 0.$$

Hence,

Corollary 1.4: $M_{\alpha\phi}$ satisfies Property 1 (Intermediate Value property).

Consistency with strict stochastic dominance ' $\overset{1}{>}$ ' (Property 3: Monotonicity) is deemed desirable for many applications of mean value. The corollary below gives the condition under which $M_{\alpha\phi}$ is consistent with stochastic dominance⁴ (nonstrict) ' $\overset{1}{\geq}$ '.

Corollary 1.5: Suppose α and ϕ are both bounded on J .

Then $\forall F, G \in D_J$, $F \overset{1}{\geq} G \Rightarrow M_{\alpha\phi}(F) \geq M_{\alpha\phi}(G)$

if and only if $\forall s \in J$, $\alpha(x)(\phi(x) - \phi(s))$ (1.26)

is a nondecreasing function (nonincreasing function).

Proof: We shall assume without loss of generality that ϕ is strictly increasing.

⁴ The partial order ' $\overset{1}{\geq}$ ' is defined by $G \overset{1}{\geq} F$ if $G(x) \leq F(x)$, $\forall x \in J$. The stronger partial order ' $\overset{1}{>}$ ' defined earlier (Property 3) is the above with strict inequality for some x .

(Sufficiency) Suppose $G \stackrel{1}{\geq} F$. Then $F_{\theta'} \stackrel{1}{\geq} F_{\theta}$ whenever $\theta' > \theta$,

$$\text{where } F_{\theta} \equiv (1-\theta)F + \theta G, \quad \forall \theta \in (0,1). \quad (1.27)$$

$$\text{Define } \zeta(x;F) \equiv \{\alpha(x)/\int_J \alpha dF\} \{\phi(x) - \Omega(F)\}, \quad (\Omega \text{ as in p. 22}) \quad (1.28)$$

$$\text{where } \Omega(F) = \int_J \alpha \phi dF / \int_J \alpha dF.$$

$$\text{Then } \frac{d}{d\theta} \Omega(F_{\theta}) = \int_J \zeta(x;F_{\theta}) d(G(x) - F(x)), \quad (1.29)$$

$$= \int_J (G(x) - F(x)) d\zeta(x;F_{\theta}) \geq 0.$$

Since the integrand is nonnegative and ζ is nondecreasing $\forall F \in D_J$,

$$\Omega(G) - \Omega(F) = \int_0^1 \left\{ \int_J (G(x) - F(x)) d\zeta(x;F_{\theta}) \right\} d\theta \geq 0$$

$$\Rightarrow M_{\alpha\phi}(G) \geq M_{\alpha\phi}(F).$$

(Necessity) Suppose $\alpha(x)(\phi(x) - \phi(s^*))$ is strictly decreasing at some x^* for some $s^* \in \text{int } J$. Since $\alpha(x)(\phi(x) - \phi(s))$ is a continuous function, it is strictly decreasing for some open neighbourhood $(x^* - \xi, x^* + \xi)$. Assume without loss of generality that $s^* \geq x^*$.

Pick any $y^* > s^*$ and compute p^* such that,

$$s^* = M_{\alpha\phi}(p^* \delta_{y^*} + (1-p^*) \delta_{x''}) = M_{\alpha\phi}(F^*), \text{ where } x'' = x^* - \frac{1}{2}\xi.$$

Consider $G^* \equiv p^* \delta_{y^*} + (1-p^*) \delta_{x'}$, for some $x' \in [x^*, x^* + \xi]$.

$$\text{Compute, } \int_J (G^*(x) - F^*(x)) d\zeta(x;F^*) = (1-p^*) (\zeta(x';F^*) - \zeta(x'';F^*)) < 0.$$

But $M_{\alpha\phi}((1-\theta)F^* + \theta G^*) = M_{\alpha\phi}(F^*_{\theta})$ is nondecreasing in θ .

$$\Rightarrow \frac{d}{d\theta} \Omega(F^*_{\theta}) = \int_J (F^*(x) - G^*(x)) d\zeta(x;F^*_{\theta}) \geq 0.$$

Since the R.H.S. is continuous, its limit as θ approaches 0 from above is nonnegative, which is a contradiction. The extension to possible end-points of J follows from the continuity and boundedness of α and ϕ .

Q.E.D.

The function $\zeta(x;F)$ can be used to generate the following linear functional,

$$\zeta_F^*(\cdot) \equiv \int_J \zeta(\cdot;F) d(\cdot) .$$

Observe that expression (1.29) is the Gâteaux differential of Ω at F_θ in the direction $G-F$, which may be written as:

$$\frac{d}{d\theta} \Omega(F_\theta) = \zeta_{F_\theta}^*(G-F) .$$

The functional $\zeta_F^*(\cdot)$ and the function $\zeta(\cdot;F)$ are both referred to as the Gâteaux derivative of Ω at F . We now interpret condition (1.26) as follows.

The Gâteaux derivative of Ω at F , $\zeta(\cdot;F)$, is nondecreasing for every F in D_J .

This generalizes the corresponding condition for quasilinear mean M_ϕ if we observe that the Gâteaux derivative of $(\phi \circ M_\phi)$ at F is simply ϕ which is strictly increasing irrespective of F .

Another useful partial order is second degree stochastic dominance ' \succeq^2 ', defined by,

$$G \succeq^2 F \text{ if } \int_J x(G(y)-F(y))dy \leq 0, \forall x \in J \text{ and } \int_J (G(y)-F(y))dy = 0$$

where $J^x = \{ y \in J : y \leq x \}$. (1.30)

The above says that G dominates F in the second degree if they have the same arithmetic mean (if they exist) and the arithmetic mean of G truncated by J^x is not less than that of F truncated by J^x for every x in J . This is equivalent to the notion of *mean preserving spread* (Rothschild & Stiglitz, 1970) in uncertainty economics, and the *principle of transfer* (Dalton, 1920) which states that a society's welfare is not diminished by a transfer of wealth from the rich to

the poor. Quasilinear mean M_ϕ is known to be consistent with second degree stochastic dominance when ϕ is increasing and concave or decreasing and convex. Having noted the similarity between ϕ and $\zeta(\cdot; F)$ in deriving consistency conditions for first degree stochastic dominance, we entertain the conjecture that the corresponding second degree condition for $M_{\alpha\phi}$ is that $\zeta(\cdot; F)$ is concave (convex) if ϕ is increasing (decreasing) for every F in D_J . The verification of this conjecture is contained as a special case of a more general result developed in the next paragraph.

We begin with the following definition of *kth degree stochastic dominance*,

$$G \stackrel{k}{\geq} F \text{ if } \int_J \{ \int_J z_{n-1} \{ \cdots \{ \int_J z_3 \{ \int_J z_2 (G(z_1) - F(z_1)) dz_1 \} dz_2 \} dz_3 \} \cdots \} dz_{n-2} \}$$

$$dz_{n-1} \} = 0,$$

for $n = 2, \dots, k$, and

$$\int_J z_k \{ \int_J z_{k-1} \{ \cdots \{ \text{as above} \} \cdots \} dz_{k-2} \} dz_{k-1} \} \leq 0, \quad \forall z_k \in J.$$

When the *nth* moment about the origin exists for distributions F and G for $n = 1, \dots, k$, then G dominates F in the *kth* degree if their *nth* moments agree for $n = 1, \dots, k$, and the *kth* moment about the origin of G truncated by J^{z_k} is not less (greater) than that of F truncated at J^{z_k} if k is odd (even) for every z_k in J . The following corollary gives conditions on α and ϕ for consistency of $M_{\alpha\phi}$ with *kth* degree stochastic dominance.

Corollary 1.6: Suppose $\alpha, \alpha', \alpha'', \dots, \alpha^{(k-1)}$, and $\phi, \phi', \phi'', \dots, \phi^{(k-1)}$ are continuous and bounded on J .

Then $\forall F, G \in D_J$, $G \stackrel{k}{\geq} F \Rightarrow M_{\alpha\phi}(G) \geq M_{\alpha\phi}(F)$

if and only if $\forall F \in D_J$, $\zeta^{(k-1)}(x; F)$

is a nondecreasing (nonincreasing) function if ϕ is increasing (decreasing) when k is odd, or

is a nonincreasing (nondecreasing) function if ϕ is increasing (decreasing) when k is even.

Proof: Assume without loss of generality that ϕ is increasing and k is even.

(Sufficiency) Suppose $G \stackrel{k}{\geq} F$. Then $F_{\theta'} \stackrel{k}{\geq} F_{\theta}$ whenever $\theta' > \theta$,

where $F \equiv (1-\theta)F + \theta G$, for $\theta \in (0,1)$.

$$\begin{aligned} \text{Then} \quad \frac{d}{d\theta} \Omega(F_{\theta}) &= \int_J \zeta(x; F_{\theta}) d(G(x) - F(x)) \\ &= (-1)^k \int_J \left\{ \int_{J^x} \{ \text{as in p. 32} \} dz_{k-1} \right\} d\zeta(x; F_{\theta}) \\ &= (-1)^k \int_J I_k(x) d\zeta(x; F_{\theta}) \\ &\geq 0, \end{aligned}$$

where $I_k(x)$ is the k -time iterated integral of $(G(x) - F(x))$ on the interval J^x (see expression 1.30),

since $I_k(x)$ is nonpositive and $\zeta^{(k-1)}$ is nondecreasing (nonincreasing) for k odd (even) $\forall F \in D_J$.

It follows that, $\Omega(G) - \Omega(F) = \int_0^1 \left\{ \int_J I_k(x) d\zeta(x; F_{\theta}) \right\} d\theta \geq 0$

$$\Rightarrow M_{\alpha\phi}(G) \geq M_{\alpha\phi}(F).$$

(Necessity) This follows from an argument that is essentially the same as the one used in the necessity proof of Corollary 1.5.

Q.E.D.

We end this section by offering a link between $M_{\alpha\phi}$ and M_{ϕ} that leads to a useful condition under which certain known inequalities

for M_ϕ can be extended to $M_{\alpha\phi}$. We derive from a distribution F , through the function α , another distribution F^α ,

$$F^\alpha(x) = \int_J \alpha dF / \int_J \alpha dF, \quad \text{for every } x \in J, \quad (1.31)$$

if the denominator exists. In this case, $M_{\alpha\phi}$ and M_ϕ are related in the following manner: $M_{\alpha\phi}(F) = M_\phi(F^\alpha)$.

This leads immediately to:

Lemma 1.3: Suppose $M_\phi(F) \geq M_\psi(F) \quad \forall F \in V \subset D_J$.

Then if $F^\alpha \in V$ whenever F does,

$$\text{then} \quad M_{\alpha\phi}(F) \geq M_{\alpha\psi}(F) \quad \forall F \in V.$$

One use of the above is the extension of the result,

$$\text{If } r > s \text{ then } M_r(F) \geq M_s(F) \text{ for every } F \in D_{(0,\infty)},$$

$$\text{where } M_r(F) = M_{x^r}(F) = \left\{ \int_0^\infty x^r dF \right\}^{1/r}.$$

to $M_{\alpha,r}(F) = \left\{ \int_0^\infty \alpha(x) x^r dF(x) / \int_0^\infty \alpha(x) dF(x) \right\}^{1/r}$. The function α has the standard measure-theoretic interpretation as a *Radon-Nikodym derivative* of F^α with respect to F . We may, on the other hand, consider F^α as an 'integral' of F through the function α . Can we define F^α even when $\int_J \alpha dF$ does not exist? Our definition of $M_{\alpha\phi}(F)$ when F does not have compact support, (expression (1.25)) suggests the following.

Let $\{K_n\}_{n=1}^\infty$ be an increasing family of compact intervals whose limit is J . Then $\int_J f dF^\alpha = \lim_{n \rightarrow \infty} \int_{K_n} \alpha f dF / \int_{K_n} \alpha dF$, for every $f \in C^0(J)$ where $C^0(J)$ denotes the space of continuous functions on J .

We have defined F^α so that the equality,

$$M_{\alpha\phi}(F) = M_\phi(F^\alpha) ,$$

holds even when F does not have compact support. It is straightforward to check that Lemma 1.3 holds for the extended definition of F^α .

GENERALIZING THE EXPECTED UTILITY REPRESENTATION THEOREM

2.1 INTRODUCTION

The preceding chapter generalized the quasilinear mean by weakening the axiom of quasilinearity. As we noted in the introduction, the quasilinear mean, M_ϕ , represents another way to axiomatize expected utility, via the maximand, $\int_R \phi dF$. Correspondingly, our generalized mean, $M_{\alpha\phi}$, induces a more general ordering via the maximand, $\int_R \alpha\phi dF / \int_R \alpha dF$.

This chapter treats the problem of extending the above representation for the case of simple probability measures on a more general outcome space than the real line. Since the notion of mean value may not be defined for a more general outcome space (consider, e.g., the outcome space consisting of getting a promotion, status quo and being fired), we need to state axioms directly in terms of properties of the underlying ordering. The developments of results parallel those of Chapter 1. Consequently, the proofs here are straightforward adaptations of the corresponding ones in Chapter 1. They are nonetheless included so that Chapter 2 may be read independently of Chapter 1. Unlike Chapter 1, most definitions used here are given explicitly because they are relatively unfamiliar.

2.2 PRELIMINARY DEFINITIONS

Definition 2.1: A simple probability measure P on a set X is a real-valued function defined on the set of all subsets of X such that:

- 1) $P(A) \geq 0$, $\forall A \subset X$;
- 2) $P(X) = 1$;
- 3) $P(A \cup B) = P(A) + P(B)$ when $A, B \subset X$ and $A \cap B = \emptyset$;
- 4) $P(A) = 1$ for some finite $A \subset X$.

A simple probability measure P on a set X has the property

$P(\{x\}) = 0$ for all but a finite number of $x \in X$ and for all

$A \subset X$, $P(A) = \sum_{x \in A} P(x)$ where $P(\{x\})$ is written as $P(x)$.

Definition 2.2: A point mass, δ_x , at x is the spm with $P(x) = 1$.

Definition 2.3: For $\beta \in (0,1)$, the β -mixture of a spm P with another

spm Q , $\beta P + (1-\beta)Q$, is the real-valued function that assigns

$\beta P(A) + (1-\beta)Q(A)$ for every $A \subset X$.

It is clear that $\beta P + (1-\beta)Q$ is a spm when P, Q are spm's. In

general, $\sum_{i=1}^n \beta_i P_i$ is a spm if P_i is a spm for $i = 1, 2, \dots, n$ and

$\sum_{i=1}^n \beta_i = 1$ with $\beta_i \geq 0$ for $i = 1, 2, \dots, n$. For a spm P on a set X ,

let $\{x_i\}_{i=1}^n \subset X$ be the set of points for which $P(x_i) > 0$ for

$i = 1, 2, \dots, n$. It is easy to check that $P \equiv \sum_{i=1}^n p_i \delta_{x_i}$, where

$p_i = P(x_i)$ for $i = 1, 2, \dots, n$.

Definition 2.4: The expectation, $E(f,P)$, of a real-valued function f

defined on X relative to a spm on X is defined by

$$E(f,P) = \sum_{x \in X} f(x)P(x)$$

For $P \equiv \sum_{i=1}^n p_i \delta_{x_i}$, $E(f,P) = \sum_{i=1}^n p_i f(x_i)$.

Definition 2.5: A binary relation $<$ on a set Y is a weak order if $<$ is

asymmetric (i.e. $\forall x, y \in Y, x < y \Rightarrow \text{not } (y < x)$) and negatively

transitive (i.e. $\forall x, y, z \in Y, \text{not } (x < y) \cap \text{not } (y < z) \Rightarrow \text{not}$

$(x < z)$).

We summarize some properties of a weak order, \prec , via the following.

Lemma 2.1: Suppose \prec is a weak order on Y . Define binary relations \sim ,

\preceq on Y by $x \sim y \Leftrightarrow \text{not } (x \prec y) \cap \text{not } (y \prec x)$, $\forall x, y \in Y$ and

$x \preceq y \Leftrightarrow (x \prec y) \cup (x \sim y)$, $\forall x, y \in Y$. Then,

- i) \sim is an equivalence relation
- ii) \preceq is transitive.
- iii) \preceq is connected (i.e. $\forall x, y \in Y$, $(x \preceq y) \cup (y \preceq x)$).
- iv) $(x \prec y) \cap (y \sim z) \Rightarrow x \prec z$, and
 $(x \sim y) \cap (y \prec z) \Rightarrow x \prec z$, $\forall x, y, z \in Y$.

Proof: (Omitted).

In a preference context, \prec is called 'strict preference' and $x \prec y$ is read as 'y is strictly preferred to x'; \preceq is called 'weak preference' and $x \preceq y$ is read as 'x is not preferred to y'; \sim is called 'indifference' and $x \sim y$ is read as 'x is indifferent to y'.

2.3 AXIOMS

The following are conditions on a binary relation \prec on L_X , the set of spm's defined on a set X .

Axiom 1: Ordering \prec is a weak order.

Axiom 2: Solvability $\forall P, Q, R \in L_X$, $P \prec Q$ and $Q \prec R$
 $\Rightarrow \exists \beta \in (0,1)$
 $\exists \beta P + (1-\beta)R \sim Q$.

Axiom 3: Monotonicity $\forall P, Q \in L_X$, $P \prec Q$
 $\Rightarrow \beta P + (1-\beta)Q \prec \gamma P + (1-\gamma)Q$
for $0 \leq \gamma < \beta \leq 1$.

Axiom 4: Weak Independence $\forall P, Q \in L_X, P \sim Q$
 $\Rightarrow \forall \beta \in (0,1) \exists \gamma \in (0,1)$
 $\ni \forall R \in L_X, \beta P + (1-\beta)R \sim \gamma Q + (1-\gamma)R.$

Axioms 1, 2, and 3 are standard properties of a binary relation that can be represented by the expectation of a utility function. Axiom 4 is our only departure. If we insist that β and γ be identical, then Axiom 4 reduces to the substitution principle, which is another property of expected utility.

The following property is a restatement of Property 5 of Chapter 1 in the context of a weak order, \preceq , on L_X

Definition 2.6: (Ratio Consistency) If $\exists P, Q, R \in L_X$
 and $\beta_1, \beta_2, \gamma_1, \gamma_2 \in (0,1) \ni$
 $P \sim Q$ and $\beta_i P + (1-\beta_i)R \sim \gamma_i Q + (1-\gamma_i)R$
 for $i = 1, 2$, then

$$\frac{\gamma_1 / 1-\gamma_1}{\beta_1 / 1-\beta_1} = \frac{\gamma_2 / 1-\gamma_2}{\beta_2 / 1-\beta_2}$$

Lemma 2.2: Axioms 1, 3 and 4 \Rightarrow Ratio Consistency.

Proof: The proof of Lemma 2.2 is essentially identical to that of Lemma 1.2 in Chapter 1.

The interpretation of our axioms and the Ratio Consistency property in the context of choice will be deferred until Chapter 4 where we apply the representation theorems of this chapter and that of Chapter 1 to decision theory.

2.4 REPRESENTATION THEOREMS

To facilitate the statement of our representation theorem, we have

Definition 2.7: Let \prec be a binary relation on L_X , the set of simple probability measures on a set X . The induced binary relation $\dot{\prec}$ on X is defined by,

$$\forall x, y \in X, x \dot{\prec} y \Leftrightarrow \delta_x \prec \delta_y.$$

If \prec is a weak order, then $\dot{\prec}$ is also a weak order. We derive the binary relations $\dot{\preceq}$ and $\dot{\succ}$ from $\dot{\prec}$ as in Lemma 2.1.

Definition 2.8: Let \prec be a weak order on a set Y , an element $y \in Y$ is a maximal (minimal) element if $\forall x \in Y, x \preceq y$ ($y \preceq x$).

Theorem 2.1: Let L_X be the set of simple probability measures defined on a set X . Suppose \prec is a binary relation on L_X with the induced binary relation on X denoted by $\dot{\prec}$. Then there exist functions $\alpha : X \rightarrow \mathbb{R}^+$ and $v : X \rightarrow \mathbb{R}$ such that v is non constant and attains its supremum and infimum over X and $\forall P, Q \in L_X$,

$$P \prec Q \Leftrightarrow \frac{E(\alpha v, P)}{E(\alpha, P)} < \frac{E(\alpha v, Q)}{E(\alpha, Q)}. \quad (2.1)$$

if and only if

\prec satisfies Axioms 1-4

and X contains a maximal element \bar{x} and a minimal element \underline{x} such that $\underline{x} \dot{\prec} \bar{x}$.

Moreover, if α, v and α^*, v^* satisfy the condition of this theorem, then $\exists a, b, c, k$ with $a, c, k > 0$ such that $\forall x \in X$,

$$\alpha^*(x) = c\alpha(x)\{k[v(x) - v(\underline{x})] + [v(\bar{x}) - v(x)]\}$$

and

$$v^*(x) = a \frac{k[v(x) - v(\underline{x})]}{k[v(x) - v(\underline{x})] + [v(\bar{x}) - v(x)]} + b. \quad (2.2)$$

Proof:

Necessity:

$$\text{Let } \underline{x}, \bar{x} \in X \ni v(\underline{x}) = \inf_{x \in X} v \text{ and } v(\bar{x}) = \sup_{x \in X} v$$

$\forall x \in X$ observe that

$$v(\underline{x}) \leq v(x) \leq v(\bar{x})$$

$$\Leftrightarrow \delta_{\underline{x}} \leq \delta_x \leq \delta_{\bar{x}}$$

$$\Leftrightarrow \underline{x} \dot{\leq} x \dot{\leq} \bar{x}$$

$\Rightarrow \underline{x}, \bar{x}$ are minimal and maximal elements of X respectively.

Furthermore, v is nonconstant $\Rightarrow \inf v < \sup v$

$$\Rightarrow \delta_{\underline{x}} < \delta_{\bar{x}}$$

$$\Leftrightarrow \underline{x} \dot{<} \bar{x}.$$

Axiom 1 follows immediately.

Axiom 2 follows from the observation that $\forall P, Q \in L_X$

$$\frac{E(\alpha v, \beta P + (1-\beta)Q)}{E(\alpha, \beta P + (1-\beta)Q)} \text{ is continuous in } \beta.$$

$$\forall P, Q \in L_X, P \dot{<} Q$$

$$\Rightarrow \frac{E(\alpha v, P)}{E(\alpha, P)} < \frac{E(\alpha v, Q)}{E(\alpha, Q)}$$

$$\Rightarrow \frac{E(\alpha v, \beta P + (1-\beta)Q)}{E(\alpha, \beta P + (1-\beta)Q)} \text{ decreases strictly in } \beta. \quad (2.3)$$

\Rightarrow Axiom 3.

Suppose $P, Q \in L_X$ and $P \sim Q \Leftrightarrow \frac{E(\alpha v, P)}{E(\alpha, P)} = \frac{E(\alpha v, Q)}{E(\alpha, Q)}$.

It follows that $\forall R \in L_X$ and $\forall \beta \in (0,1)$

$$\frac{E(\alpha v, \beta P + (1-\beta)R)}{E(\alpha, \beta P + (1-\beta)R)} = \frac{E(\alpha v, \gamma Q + (1-\gamma)R)}{E(\alpha, \gamma Q + (1-\gamma)R)}$$

$$\text{where } \frac{\gamma/(1-\gamma)}{\beta/(1-\beta)} = \frac{E(\alpha, P)}{E(\alpha, Q)}.$$

Hence, Axiom 4.

Sufficiency:

Let \underline{x} , \bar{x} be minimal and maximal elements of X , respectively.

Define $\forall p \in [0,1]$,

$$S_p \equiv p\delta_{\bar{x}} + (1-p)\delta_{\underline{x}}.$$

By hypothesis $\delta_{\underline{x}} \prec \delta_{\bar{x}}$. It follows from Axiom 3 that

$$S_p \prec S_q \Leftrightarrow 0 \leq p < q \leq 1. \quad (2.4)$$

$\forall x \in X \ni \delta_{\underline{x}} \prec \delta_x$ and $\delta_x \prec \delta_{\bar{x}}$, it follows from Axiom 2 that

$$\exists q \in (0,1) \ni \delta_x \sim S_q.$$

It is clear from (2.4) that q is unique.

We construct a real-valued function $\tilde{v} : X \rightarrow [0,1]$ in the following manner.

$$\tilde{v}(\underline{x}) = 0, \quad \tilde{v}(\bar{x}) = 1,$$

$$\forall x \in X - \{\underline{x}, \bar{x}\}, \quad \tilde{v}(x) = q \ni \delta_x \sim S_q.$$

From construction, $\forall x \in X - \{\underline{x}, \bar{x}\}$,

$$\delta_x \sim S_{\tilde{v}(x)} \quad (2.5)$$

Lemma 2.2 $\Rightarrow \exists \tau_x > 0 \ni \forall R \in L_X$ and $\beta \in (0,1)$

$$\beta \delta_x + (1-\beta)R \sim \frac{\beta \tau_x}{\beta \tau_x + 1 - \beta} S_{\tilde{v}(x)} + \frac{1-\beta}{\beta \tau_x + 1 - \beta} R \quad (2.6)$$

Construct a positive real-valued function $\tilde{\alpha}$ on X in the following manner.

$$\tilde{\alpha}(\underline{x}) = \tilde{\alpha}(\bar{x}) = 1$$

$$\forall x \in X - \{\underline{x}, \bar{x}\}, \tilde{\alpha}(x) = \tau_x.$$

Given a spm $P \equiv \sum_{i=1}^n p_i \delta_{x_i}$.

Applying (2.6) to x_1, x_2, \dots, x_n sequentially, it follows that

$$\begin{aligned} \sum_{i=1}^n p_i \delta_{x_i} &\sim \frac{p_1 \tilde{\alpha}(x_1)}{p_1 \tilde{\alpha}(x_1) + \sum_{j=2}^n p_j} S_{\tilde{v}(x_1)} + \sum_{i=2}^n \frac{p_i}{p_1 \tilde{\alpha}(x_1) + \sum_{j=2}^n p_j} \delta_{x_i}, \\ &\sim \frac{p_1 \tilde{\alpha}(x_1)}{p_1 \tilde{\alpha}(x_1) + p_2 \tilde{\alpha}(x_2) + \sum_{j=3}^n p_j} S_{\tilde{v}(x_1)} \\ &\quad + \frac{p_2 \tilde{\alpha}(x_2)}{p_1 \tilde{\alpha}(x_1) + p_2 \tilde{\alpha}(x_2) + \sum_{j=3}^n p_j} S_{\tilde{v}(x_2)} \\ &\quad + \sum_{i=3}^n \frac{p_i}{p_1 \tilde{\alpha}(x_1) + p_2 \tilde{\alpha}(x_2) + \sum_{j=3}^n p_j} \delta_{x_i} \\ &\quad \vdots \end{aligned}$$

$$\sim \sum_{i=1}^n \frac{p_i \tilde{\alpha}(x_i)}{\sum_{j=1}^n p_j \tilde{\alpha}(x_j)} S_{\tilde{v}(x_i)} \equiv S_{\left[\sum_{i=1}^n p_i \tilde{\alpha}(x_i) \tilde{v}(x_i) / \sum_{j=1}^n p_j \tilde{\alpha}(x_j) \right]}.$$

Hence, $\forall P \in L_X, P \sim \frac{S_{E(\tilde{\alpha}\tilde{v}, P)}}{E(\tilde{\alpha}, P)}.$

It follows from Lemma 2.1 (iv) that $\forall P, Q \in L_X,$

$$\begin{aligned} P \prec Q &\Leftrightarrow \frac{S_{E(\tilde{\alpha}\tilde{v}, P)}}{E(\tilde{\alpha}, P)} \prec \frac{S_{E(\tilde{\alpha}\tilde{v}, Q)}}{E(\tilde{\alpha}, Q)} \\ &\Leftrightarrow \frac{E(\tilde{\alpha}\tilde{v}, P)}{E(\tilde{\alpha}, P)} < \frac{E(\tilde{\alpha}\tilde{v}, Q)}{E(\tilde{\alpha}, Q)}. \end{aligned}$$

Uniqueness:

Suppose $\exists \alpha : X \rightarrow R^+$ and $v : X \rightarrow R$ such that v is nonconstant and attains its infimum and supremum over X at \underline{y}, \bar{y} , respectively, and $\forall P, Q \in L_X,$

$$P \prec Q \Leftrightarrow \frac{E(\alpha v, P)}{E(\alpha, P)} < \frac{E(\alpha v, Q)}{E(\alpha, Q)} \quad (2.7)$$

Clearly, \underline{y}, \bar{y} are minimal and maximal elements of X .

$$\Rightarrow (\underline{y} \equiv \underline{x}) \cup (\delta_{\underline{y}} \sim \delta_{\underline{x}}) \text{ and } (\bar{y} \equiv \bar{x}) \cup (\delta_{\bar{y}} \sim \delta_{\bar{x}}).$$

$$\Rightarrow v(\underline{x}) = v(\underline{y}) \text{ and } v(\bar{x}) = v(\bar{y}).$$

By construction, $\forall x \in X - \{\underline{x}, \bar{x}\},$

$$\delta_x \sim S_{\tilde{v}(x)}, \quad (\text{see relation (2.5)}) \quad (2.8)$$

$$\Leftrightarrow v(x) = \frac{\tilde{v}(x)\alpha(\bar{x})v(\bar{x}) + (1 - \tilde{v}(x))\alpha(\underline{x})v(\underline{x})}{\tilde{v}(x)\alpha(\bar{x}) + (1 - \tilde{v}(x))\alpha(\underline{x})}, \quad (2.9)$$

after applying (2.7) to (2.8).

Also $\forall \beta \in (0,1)$,

$$\beta \delta_{\underline{x}} + (1-\beta) \delta_{\underline{x}} \sim \frac{\beta \tilde{\alpha}(x) S_{\tilde{v}(x)} + (1-\beta) \delta_{\underline{x}}}{\beta \tilde{\alpha}(x) + (1-\beta)} \quad (2.10)$$

$$\Leftrightarrow \frac{\beta \alpha(x) v(x) + (1-\beta) \alpha(\underline{x}) v(\underline{x})}{\beta \alpha(x) + (1-\beta) \alpha(\underline{x})} \quad (2.11)$$

$$= \frac{\beta \tilde{\alpha}(x) \tilde{v}(x) \alpha(\bar{x}) v(\bar{x}) + [\beta \tilde{\alpha}(x) (1-\tilde{v}(x)) + (1-\beta)] \alpha(\underline{x}) v(\underline{x})}{\beta \tilde{\alpha}(x) \tilde{v}(x) \alpha(\bar{x}) + [\beta \tilde{\alpha}(x) (1-\tilde{v}(x)) + (1-\beta)] \alpha(\underline{x})}$$

after applying (2.7) to (2.10)

Let $a = v(\bar{x}) - v(\underline{x})$,

$b = v(\underline{x})$,

$c = \alpha(\underline{x})$,

$k = \frac{\alpha(\bar{x})}{\alpha(\underline{x})}$.

It is easy to check that (2.9) and (2.11) become

$$v(x) = a \frac{k \tilde{v}(x)}{k \tilde{v}(x) + (1 - \tilde{v}(x))} + b$$

$$\text{and} \quad \alpha(x) = c \tilde{\alpha}(x) [k \tilde{v}(x) + (1 - \tilde{v}(x))]. \quad (2.12)$$

Suppose α^* , v^* are another pair of functions that satisfy the hypotheses of the theorem. Then

$$v^*(x) = a^* \frac{k^* \tilde{v}(x)}{k^* \tilde{v}(x) + (1 - \tilde{v}(x))} + b^*$$

$$\text{and} \quad \alpha^*(x) = c^* \tilde{\alpha}(x) [k^* \tilde{v}(x) + (1 - \tilde{v}(x))], \quad (2.13)$$

where $a^* = v^*(\bar{x}) - v^*(\underline{x})$, $b^* = v^*(\underline{x})$,

$$c^* = \alpha^*(\underline{x}) \text{ , and } k^* = \frac{\alpha^*(\bar{x})}{\alpha^*(\underline{x})} .$$

Finally, it is straightforward to check that,

$$v^*(x) = a' \frac{k' [v(x) - v(\underline{x})]}{k' [v(x) - v(\underline{x})] + v(\bar{x}) - v(x)} + b'$$

$$\text{and } \alpha^*(x) = c' \alpha(x) \{k' [v(x) - v(\underline{x})] + v(\bar{x}) - v(x)\} \quad (2.14)$$

$$\text{for } a' = (v^*(\bar{x}) - v^*(\underline{x})) / a = a^* / a$$

$$b' = v^*(\underline{x}) = b^*$$

$$c' = \frac{\alpha^*(\underline{x})}{\alpha(\underline{x})} \left(\frac{1}{v(\bar{x}) - v(\underline{x})} \right) = \frac{c^*}{c} \cdot \frac{1}{a}$$

$$k' = \frac{\alpha^*(\bar{x})}{\alpha(\bar{x})} \frac{\alpha(\underline{x})}{\alpha^*(\underline{x})} = \frac{k^*}{k} .$$

Q.E.D.

We showed that any binary relation on L_X , that satisfies the ordering, monotonicity, solvability and weak independence axioms, is characterized by a pair of functions (α, v) defined on X . When α is constant, our representation, $E(\alpha v, P) / E(\alpha, P)$, reduces to the expected utility representation $E(v, P)$. If we define a simple probability measure P^α derived from P in the following manner,

$$P^\alpha(A) = E(\alpha 1_A, P) / E(\alpha, P) \text{ , } \forall A \subset X,$$

where 1_A denotes the indicator function of A ,

then we can state our representation in the alternative fashion below:

$$E(v, P^\alpha) .$$

Our representation is then simply the expectation of the v -function

with respect to the measure P^α derived from P via the α -function which is the Radon-Nikodym derivative of P^α with respect to P . We render the role of α transparent by considering a uniform measure

$P \equiv \sum_{i=1}^N \frac{1}{N} \delta_{x_i}$. The corresponding representation of P is given by,

$$E(\alpha v, P) / E(\alpha, P) = \sum_{i=1}^N \frac{\alpha(x_i)}{\sum_{j=1}^N \alpha(x_j)} v(x_i).$$

This is a weighted average of $\{v(x_i)\}_{i=1}^N$ with weights, $\{\alpha(x_i)\}_{i=1}^N$.

The statement of Theorem 2.1 requires the set X to be bounded by a maximal and a minimal element. The remainder of this section deals with the extension of Theorem 2.1 to the case where X has neither a maximal nor a minimal element. This parallels the development towards the proof of Theorem 1.3 in Chapter 1.

Definition 2.9: Let $<$ be a weak order on a set X . For $s, t \in X$, an interval $[s, t] \subset X$ is defined by $[s, t] = \{x \in X : s \leq x, x \leq t\}$.

When X contains both a maximal element \bar{x} and a minimal element \underline{x} relative to a weak order $<$, then $X = [\underline{x}, \bar{x}]$. When $s_1, s_2, t_1, t_2 \in X$, such that $s_2 < s_1 < t_1 < t_2$, then $[s_1, t_1] \subsetneq [s_2, t_2]$.

Definition 2.10: A pair of functions (α, v) is said to represent a weak order $<$ on L_X if (α, v) satisfies condition (2.1).

Definition 2.11: The uniqueness class representing a weak order \prec on L_X , $\{\alpha, \nu\}_X$, consists of all pairs (α, ν) that represent \prec on L_X .

Definition 2.12: Let $s, t \in X \ni s \prec t$, we denote by $\{\alpha, \nu\}_{[s,t]}^k$, the k-ratio subclass of the uniqueness class $\{\alpha, \nu\}_{[s,t]}$ representing \prec on $L_{[s,t]}$, consisting of those pairs (α, ν) that satisfy

$$\frac{\alpha(t)}{\alpha(s)} = k. \text{ A generic element of } \{\alpha, \nu\}_{[s,t]}^k \text{ is denoted by } (\alpha^k, \nu^k).$$

Definition 2.13: Let $s, t \in X \ni s \prec t$, the pair (α^*, ν^*) is said to be an (a,b,c,k) transformation of (α, ν) on $[s,t]$ if $\exists a, b, c, k$ such that $a, c, k > 0$ and $\forall x \in [s,t]$

$$\alpha^*(x) = c\alpha(x)\{k(\nu(x) - \nu(s)) + \nu(t) - \nu(x)\}$$

$$\text{and } \nu^*(x) = a \frac{k[\nu(x) - \nu(s)]}{k[\nu(x) - \nu(s)] + \nu(t) - \nu(x)} + b.$$

We denote such a transformation by

$$(\alpha^*, \nu^*) = T_{a,b,c,k}(\alpha, \nu) \text{ on } [s,t] \subset X.$$

It is clear that $\{\alpha, \nu\}_{[s,t]} = \bigcup_{k \in \mathbb{R}^+} \{\alpha, \nu\}_{[s,t]}^k$. Let (α, ν) be an element of $\{\alpha, \nu\}_{[s,t]}$, we can then generate all other elements via the (a,b,c,k) transformation. Note that $T_{a,b,1,1}$ is an affine transformation on ν and $T_{1,1,c,1}$ is a positive scalar multiple of α ; and that we can use a unitary ratio pair

$$(\alpha^1, \nu^1) \in \{\alpha, \nu\}_{[s,t]}^1 \text{ to generate all elements of } \{\alpha, \nu\}_{[s,t]}$$

since $T_{a,b,c,k}(\alpha^1, \nu^1) \in \{\alpha, \nu\}_{[s,t]}^k$. We denote by $(\tilde{\alpha}^k, \tilde{\nu}^k)$, the canonical member of $\{\alpha, \nu\}_{[s,t]}^k$ which satisfies $\tilde{\nu}^k(s) = 0$, $\tilde{\nu}^k(t) = 1$, $\tilde{\alpha}^k(s) = 1$, and $\tilde{\alpha}^k(t) = k$. It is clear that $(\tilde{\alpha}^k, \tilde{\nu}^k)$ is unique for each k . In general, a member α, ν of $\{\alpha, \nu\}_{[s,t]}$ is uniquely specified by the values of α, ν at s and t . Let $a = \nu(t) - \nu(s)$, $b = \nu(s)$, $c = \alpha(s)$, $k = \alpha(t)/\alpha(s)$, then $\forall x \in [s, t]$.

$$\alpha(x) = c\tilde{\alpha}^1(x)[k\tilde{\nu}^1(x) + (1 - \tilde{\nu}^1(x))]$$

and

$$\nu(x) = a \frac{k\tilde{\nu}^1(x)}{k\tilde{\nu}^1(x) + 1 - \tilde{\nu}^1(x)} + b.$$

Corollary 2.1: Let $s_0, s_1, t_0, t_1 \in X \ni s_1 \prec s_0 \prec t_0 \prec t_1$, then

$$\{\alpha, \nu\}_{[s_1, t_1]}|_{[s_0, t_0]} = \bigcup_{k \in (\underline{l}_{01}, \bar{l}_{01})} \{\alpha, \nu\}_{[s_0, t_0]}^k, \text{ where}$$

$$\bar{l}_{01} = \frac{\tilde{\alpha}^1(t_0)}{\tilde{\alpha}^1(s_0)} \frac{\tilde{\nu}^1(t_0)}{\tilde{\nu}^1(s_0)}, \quad (2.15)$$

$$\underline{l}_{01} = \frac{\tilde{\alpha}^1(t_0)}{\tilde{\alpha}^1(s_0)} \frac{1 - \tilde{\nu}^1(t_0)}{1 - \tilde{\nu}^1(s_0)}. \quad (2.16)$$

Proof: $[s_0, t_0] \subsetneq [s_1, t_1] \Rightarrow L_{[s_0, t_0]} \subsetneq L_{[s_1, t_1]}$. Therefore, if α, ν

defined on $[s_1, t_1]$ represents \prec on $L_{[s_1, t_1]}$, then $\alpha, \nu|_{[s_0, t_0]}$

represents $<$ on $L_{[s_0, t_0]}$, $\Rightarrow \{\alpha, \nu\}_{[s_1, t_1]} \Big|_{[s_0, t_0]} \subset \{\alpha, \nu\}_{[s_0, t_0]}$.

Observe that $\{\alpha, \nu\}_{[s_1, t_1]}^{k_1} \Big|_{[s_0, t_0]} = \{\alpha, \nu\}_{[s_0, t_0]}^{k_0}$, where

$$k_0 = \frac{\tilde{\alpha}^1(t_0)}{\tilde{\alpha}^1(s_0)} \frac{k_1 \tilde{\nu}^1(t_0) + (1 - \tilde{\nu}^1(t_0))}{k_1 \tilde{\nu}^1(s_0) + (1 - \tilde{\nu}^1(s_0))} \quad (2.17)$$

$$= \xi(k_1).$$

Note that ξ is a continuous, strictly increasing and onto function from $(0, \infty)$ to $(\underline{\ell}_{01}, \bar{\ell}_{01})$. Hence

$$\begin{aligned} \{\alpha, \nu\}_{[s_1, t_1]} \Big|_{[s_0, t_0]} &= \bigcup_{k_1 \in (0, \infty)} \{\alpha, \nu\}_{[s_1, t_1]}^{k_1} \Big|_{[s_0, t_0]} \\ &= \bigcup_{k_1 \in (0, \infty)} \{\alpha, \nu\}_{[s_0, t_0]}^{\xi(k_1)} \\ &= \bigcup_{k_0 \in (\underline{\ell}_{01}, \bar{\ell}_{01})} \{\alpha, \nu\}_{[s_0, t_0]}^{k_0}. \end{aligned}$$

Q.E.D.

Corollary 2.1 tells us it is possible to extend our representation of a weak order on $L_{[s_0, t_0]}$ to a weak order on $L_{[s_1, t_1]}$ if the ordering has not changed on $L_{[s_0, t_0]}$. It also gives conditions on $(\alpha_0, \nu_0) \in \{\alpha, \nu\}_{[s_0, t_0]}$ so that (α_0, ν_0) can be extended to a

member $(\alpha_1, \nu_1) \in \{\alpha, \nu\}_{[s_1, t_1]}$ such that $(\alpha_0, \nu_0) \equiv (\alpha_1, \nu_1) \Big|_{[s_0, t_0]}$,

i.e. the extended pair (α_1, ν_1) defined on $[s_1, t_1]$ agrees with

(α_0, ν_0) on $[s_0, t_0]$. These conditions ((2.15) and (2.16)) are given

in terms of $\tilde{\alpha}^1(t_0)$, $\tilde{\alpha}^1(s_0)$, $\tilde{\nu}^1(t_0)$, $\tilde{\nu}^1(s_0)$; and can be obtained

via relations (2.5) and (2.6) in the constructive proof of

Theorem 2.1.

Corollary 2.2: Let $s_0, s_1, s_2, t_0, t_1, t_2 \in X \ni s_2 \dot{\prec} s_1 \dot{\prec} s_0 \dot{\prec} t_0$

$\dot{\prec} t_1 \dot{\prec} t_2$, and let $(\tilde{\alpha}_i^1, \tilde{\nu}_i^1)$ denote the canonical unitary α, ν

on $[s_i, t_i]$, for $i = 1, 2$. Then,

$$\bar{\ell}_{02} < \bar{\ell}_{01} \text{ and } \ell_{-02} > \ell_{-01} \quad (2.18)$$

$$\text{where } \bar{\ell}_{0i} = \frac{\tilde{\alpha}_i^1(t_0)}{\tilde{\alpha}_i^1(s_0)} \frac{\tilde{\nu}_i^1(t_0)}{\tilde{\nu}_i^1(s_0)}$$

$$\text{and } \ell_{-0i} = \frac{\tilde{\alpha}_i^1(t_0)}{\tilde{\alpha}_i^1(s_0)} \frac{1 - \tilde{\nu}_i^1(t_0)}{1 - \tilde{\nu}_i^1(s_0)}, \quad \text{for } i = 1, 2.$$

Proof: From Corollary 2.1,

$$\{\alpha, \nu\}_{[s_2, t_2]} \Big|_{[s_0, t_0]} = \bigcup_{k_0 \in (\ell_{-02}, \bar{\ell}_{02})} \{\alpha, \nu\}_{[s_0, t_0]}^{k_0},$$

$$\{\alpha, \nu\}_{[s_1, t_1]} \Big|_{[s_0, t_0]} = \bigcup_{k_0 \in (\ell_{-01}, \bar{\ell}_{01})} \{\alpha, \nu\}_{[s_0, t_0]}^{k_0},$$

and
$$\{\alpha, \nu\}_{[s_2, t_2]} \Big|_{[s_1, t_1]} = \bigcup_{k_1 \in (\ell_{12}, \bar{\ell}_{12})} \{\alpha, \nu\}_{[s_1, t_1]}^{k_1}$$

where
$$\bar{\ell}_{12} = \frac{\tilde{\alpha}_2^1(t_1)}{\tilde{\alpha}_2^1(s_1)} \frac{\tilde{\nu}_2^1(t_1)}{\tilde{\nu}_2^1(s_1)},$$

and
$$\ell_{12} = \frac{\tilde{\alpha}_2^1(t_1)}{\tilde{\alpha}_2^1(s_1)} \frac{1 - \tilde{\nu}_2^1(t_1)}{1 - \tilde{\nu}_2^1(s_1)}.$$

Construct the function $\xi_{j,i} : (0, \infty) \rightarrow (\ell_{ij}, \bar{\ell}_{ij})$, for $i < j$, via

$$\begin{aligned} k_i &= \xi_{j,i}(k_j) \\ &= \frac{\tilde{\alpha}_j^1(t_i)}{\tilde{\alpha}_j^1(s_i)} \frac{k_j \tilde{\nu}_j^1(t_i) + 1 - \tilde{\nu}_j^1(t_i)}{k_j \tilde{\nu}_j^1(s_i) + 1 - \tilde{\nu}_j^1(s_i)}. \end{aligned} \quad (2.19)$$

Note that, by construction,

$$\{\alpha, \nu\}_{[s_j, t_j]}^{k_j} \Big|_{[s_i, t_i]} = \{\alpha, \nu\}_{[s_i, t_i]}^{\xi_{j,i}(k_j)}.$$

Note also that $\xi_{j,i}$ is continuous, strictly increasing, and onto from $(0, \infty)$ to $(\ell_{ij}, \bar{\ell}_{ij})$.

Suppose $\bar{\ell}_{02} \geq \bar{\ell}_{01}$.

Pick $k_1 \in [\bar{\ell}_{12}, \infty)$.

then $\xi_{10}(k_1) < \bar{\ell}_{01} \leq \bar{\ell}_{02}$.

$\Rightarrow \exists k_2 \in (0, \infty) \ni \xi_{20}(k_2) = \xi_{10}(k_1)$.

$$\text{i.e. } \{\alpha, \nu\}_{[s_2, t_2]}^{k_2} \Big|_{[s_0, t_0]} = \{\alpha, \nu\}_{[s_0, t_0]}^{\xi_{10}(k_1)}.$$

$$\begin{aligned} \text{But } & \{\alpha, \nu\}_{[s_2, t_2]}^{k_2} \Big|_{[s_0, t_0]} \\ &= \{\alpha, \nu\}_{[s_2, t_2]}^{k_2} \Big|_{[s_1, t_1]} \Big|_{[s_0, t_0]} \\ &= \{\alpha, \nu\}_{[s_1, t_1]}^{\xi_{21}(k_2)} \Big|_{[s_0, t_0]} \\ &\Rightarrow \xi_{21}(k_2) = k_1 \end{aligned}$$

$$\text{But } \xi_{21}(k_2) \in (\underline{\ell}_{12}, \bar{\ell}_{12}) \neq .$$

A similar argument establishes $\underline{\ell}_{01} < \underline{\ell}_{02}$.

Q.E.D.

Thus far, we have considered extending from an interval of X to a larger interval. Presumably X is not bounded, otherwise, we would have constructed (α, ν) on X with Theorem 2.1 once and for all. The next interesting case then is when X contains neither a maximal nor a minimal element, for example, the real line. With a structural condition on X , we show in Theorem 2.2 that even in this case, a α - ν representation exists on L_X .

Definition 2.14: Let \preceq be a weak order on a set Y . A sequence

$\{y_i\}_{i=1}^{\infty} \subset Y$ is cofinal (coinitial) if $\forall x \in Y, x \preceq y_i (y_i \preceq x)$ for some positive integer i .

Theorem 2.2: Let L_X be the set of spm defined on a set X , and \prec is a binary relation on L_X with the induced binary relation on X denoted by \prec . There exist functions $\alpha : X \rightarrow \mathbb{R}^+$ and $v : X \rightarrow \mathbb{R}$ such that

- (i) $v(X)$ contains a strictly increasing cofinal sequence and a strictly decreasing coinital sequence, and
- (ii) $\forall P, Q \in L_X$

$$P \prec Q \Leftrightarrow \frac{E(\alpha v, P)}{E(\alpha, P)} < \frac{E(\alpha v, Q)}{E(\alpha, Q)},$$

if and only if,

\prec satisfies Axioms 1-5

and X ordered by \prec contains a strictly increasing cofinal sequence and a strictly decreasing coinital sequence.

Moreover, if α, v and α^*, v^* both satisfy (i) and (ii), then

$\forall s, t \in X \ni s \prec t, \exists a, b, c, k$ with $a, c, k > 0 \ni \forall x \in [s, t],$

$$\alpha^*(x) = c\alpha(x)\{k[v(x) - v(s)] + [v(t) - v(x)]\},$$

$$\text{and } v^*(x) = a \frac{k[v(x) - v(s)]}{k[v(x) - v(s)] + [v(t) - v(x)]} + b.$$

Proof:

Necessity:

Let $\{d_i\}_{i=0}^{\infty}$ and $\{e_i\}_{i=0}^{\infty} \subset v(X)$ be a strictly decreasing coinital sequence and a strictly increasing cofinal sequence, respectively.

Pick $s_i, t_i \in X \ni v(s_i) = d_i$ and $v(t_i) = e_i$ for $i = 0, 1, 2, \dots$

It follows that $\{s_i\}_{i=1}^{\infty}$ ($\{t_i\}_{i=1}^{\infty}$) is a strictly decreasing
 cointial sequence (strictly increasing cofinal sequence) of χ .
 Verification of Axioms 1-4 is straightforward (see Necessity
 Proof of Theorem 2.1).

Sufficiency:

Let $\{s_i\}_{i=0}^{\infty}$ ($\{t_i\}_{i=0}^{\infty}$) be a strictly decreasing cointial sequence
 (strictly increasing cofinal sequence) of χ . Suppose without
 loss of generality that $s_0 \prec t_0$. It is easy to check that

$$\chi = \bigcup_{i=0}^{\infty} [s_i, t_i].$$

Let (α_i, ν_i) represent \prec on $L_{[s_i, t_i]}$, for $i = 0, 1, 2, \dots$. It can
 always be done because of Theorem 2.1.

Corollary 2.1 \Rightarrow

$$\{\alpha_i, \nu_i\}_{[s_i, t_i]} \big|_{[s_0, t_0]} = \bigcup_{k_0 \in (\ell_{0i}, \bar{\ell}_{0i})} \{\alpha_0, \nu_0\}_{[s_0, t_0]}^{k_0},$$

$$\text{where } (\ell_{0i}, \bar{\ell}_{0i}) = \left(\frac{\tilde{\alpha}_i^1(t_0)}{\tilde{\alpha}_i^1(s_0)} \cdot \frac{1 - \tilde{\nu}_i^1(t_0)}{1 - \tilde{\nu}_i^1(s_0)}, \frac{\tilde{\alpha}_i^1(t_0)}{\tilde{\alpha}_i^1(s_0)} \cdot \frac{\tilde{\nu}_i^1(t_0)}{\tilde{\nu}_i^1(s_0)} \right).$$

Corollary 2.2 \Rightarrow

$\{\ell_{0i}\}_{i=1}^{\infty}$ is strictly increasing and $\{\bar{\ell}_{0i}\}_{i=1}^{\infty}$ is
 strictly decreasing.

$$\text{Let } A_i = \left[\frac{\ell_{0i} + \ell_{0,i+1}}{2}, \frac{\bar{\ell}_{0i} + \bar{\ell}_{0,i+1}}{2} \right],$$

$$\text{and } B_i = (\ell_{0,i+1}, \bar{\ell}_{0,i+1}).$$

Then $B_i \subsetneq A_i \subsetneq (\ell_{0i}, \bar{\ell}_{0i})$ for $i = 1, 2, 3, \dots$.

Observe that $(\ell_{0i}, \bar{\ell}_{0i})$, A_i , B_i are strictly decreasing sequences by inclusion. Since A_i is compact for each i ,

$$\text{nested interval theorem} \Rightarrow \lim_{i \rightarrow \infty} A_i = A_\infty \neq \emptyset.$$

Since $B_i \subsetneq A_i \subsetneq (\ell_{0i}, \bar{\ell}_{0i}) \quad \forall i$

$$\Rightarrow \lim_{i \rightarrow \infty} B_i \subset A_\infty \subset \lim_{i \rightarrow \infty} (\ell_{0i}, \bar{\ell}_{0i}).$$

But $\lim_{i \rightarrow \infty} B_i = \lim_{i \rightarrow \infty} (\ell_{0i}, \bar{\ell}_{0i})$.

Hence $\lim_{i \rightarrow \infty} (\ell_{0i}, \bar{\ell}_{0i}) = A_\infty$.

To construct (α, ν) defined on X that represents \prec on L_X , pick $k_0 \in A_\infty$.

Define $(\alpha(x), \nu(x)) = (\tilde{\alpha}_0^{k_0}(x), \tilde{\nu}_0^{k_0}(x))$ for $x \in [s_0, t_0]$.

$$= (\alpha_1^{k_1}(x), \nu_1^{k_1}(x)) \text{ for } x \in [s_1, t_1] - [s_0, t_0],$$

⋮

$$= (\alpha_i^{k_i}(x), \nu_i^{k_i}(x)) \text{ for } x \in [s_i, t_i] - [s_{i-1}, t_{i-1}],$$

⋮

such that $\xi_{i0}(k_i) = k_0 = \tilde{\alpha}_0^{k_0}(t_0)$ (see (2.19)),

and $\alpha_i^{k_i}(s_0) = 1 = \tilde{\alpha}_0^{k_0}(s_0)$.

$$\nu_i^{k_i}(s_0) = 0 = \tilde{\nu}_0^{k_0}(s_0).$$

$$\nu_i^{k_i}(t_0) = 1 = \tilde{\nu}_0^{k_0}(t_0).$$

Observe that $(\alpha_{i+1}^{k_{i+1}}, v_{i+1}^{k_{i+1}}) \Big|_{[s_i, t_i]} \equiv (\alpha_i^{k_i}, v_i^{k_i})$ since they

agree at s_0 and at t_0 .

$\forall P, Q \in L_X$, pick $[s_i, t_i]$

$\ni P, Q \in L_{[s_i, t_i]}$,

$$\text{then } P < Q \Leftrightarrow \frac{E(\alpha_i^{k_i}, v_i^{k_i}, P)}{E(\alpha_i^{k_i}, P)} < \frac{E(\alpha_i^{k_i}, v_i^{k_i}, Q)}{E(\alpha_i^{k_i}, Q)}$$

$$\Leftrightarrow \frac{E(\alpha v, P)}{E(\alpha, P)} < \frac{E(\alpha v, Q)}{E(\alpha, Q)}$$

by construction.

To complete the sufficiency proof, observe that $\{v(s_i)\}_{i=0}^{\infty}$

$(\{v(t_i)\}_{i=0}^{\infty})$ is a strictly decreasing coinital sequence

(strictly increasing cofinal sequence) of $v(X)$.

Uniqueness:

This follows directly from applying Theorem 2.1 to arbitrary intervals $[s, t]$ in X .

Q.E.D.

PART II

APPLICATION TO DECISION THEORY

BACKGROUND

A choice situation exists when more than one course of action is available to a decision maker. A theory of choice specifies, for each set of available alternatives, the one that will be chosen. We have a valid descriptive theory if, for the relevant domain of choice situations, the theory can be compatible with the actual choices. The theory is normatively compelling if the underlying postulates are of sufficient appeal so that a decision maker is willing to change his choice to conform to the theory's specifications. Expected utility has been considered an example of such a theory because for many researchers (e.g. Savage (1954), MacCrimmon (1965) and Raiffa (1968)), it satisfies the latter requirements. Yet, there is enough empirical evidence (cf. Chapter 3) to suggest that it is not a very good descriptive theory. People tend to systematically violate the implications of a key property of expected utility called the strong independence principle or the substitution axiom. Many of them would not change their choices after being told of their violations (MacCrimmon, 1968; Slovic & Tversky, 1975).

Due to the success of expected utility in the modeling of phenomena in the economics of uncertainty and its application to statistical decision theory, it has been fashionable to discount violations as mistakes needing correction. A departure from this trend is evident in the appearance of several recent papers (Meginniss, 1977; Handa, 1977; Karmarkar, 1978; Kahneman & Tversky, 1979; Machina, 1980) proposing alternative theories of choice to account for Allais' 'paradox' and

other empirical findings that contradict the implications of expected utility.

We develop in Chapter 4 a new theory of choice called alpha utility theory which generalizes expected utility via a necessary and sufficient set of axioms that weaken the corresponding ones for expected utility. Specifically, the substitution axiom is replaced by a weaker axiom called *Weak Independence*. Given two lotteries that are indifferent to each other, Weak Independence allows for different probabilities in composing each of these lotteries with a third lottery to preserve indifference. However, these mixture-probabilities once determined must be independent of the third lottery. The axioms imply that the ratio of the mixture (probability) odds is constant. We call this the *Ratio Consistency* property. Expected utility results when this ratio is identically unity. Our theory has descriptive relevance in that it can represent the usual responses given to the Allais paradox and is compatible with other reported empirical findings contradicting the implications of expected utility. Yet, it can be consistent with such normatively appealing partial orders as stochastic dominance and global risk aversion. As with expected utility, the constructive proof of our representation theorem furnishes a procedure for the assessment of the alpha utility functions, from which empirically testable predictions can be derived.

CRITIQUE OF EXPECTED UTILITY THEORY

3.1 INTRODUCTION

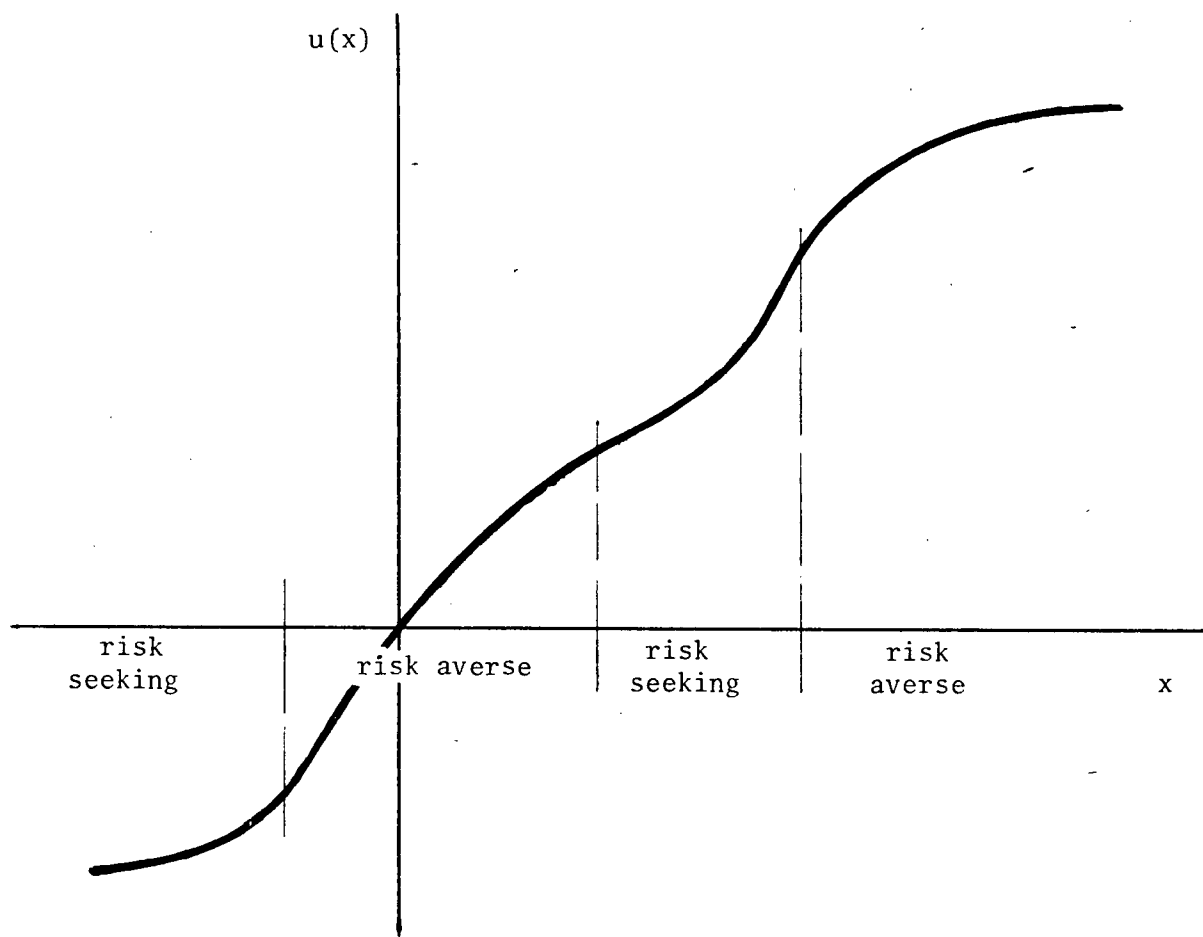
Expected utility theory has attracted considerable attention since its revival by von Neumann and Morgenstern in their "Theory of Games and Economic Behavior". It serves as the foundation for the economics of uncertainty (Arrow, 1971; Marschak and Radner, 1972; Diamond and Rothschild, 1978), statistical decision theory (Savage, 1954; Blackwell and Girshick, 1961; Raiffa and Schlaifer, 1961; DeGroot, 1970) and decision analysis (Howard, 1964; Keeney and Raiffa, 1976). Expected utility has been less successful though in explaining and describing actual choices (Edwards, 1954, 1961; MacCrimmon, 1965; Kahneman and Tversky, 1979); thus, providing a strong impetus for further theoretical development.

In this critique, we review briefly the empirical findings that pose difficulty for expected utility. First, we provide a summary, based on a recent paper (Chew and MacCrimmon, 1979b), of the systematic violations of the strong independence principle. The first example of such a violation is provided by the Allais paradox, which inspired extensive follow-up studies and modifications.

The next section, discusses the concurrence of risk-averting and risk-seeking behavior evident in the prevalence of the purchase of insurance and gambling (Friedman and Savage, 1948). For instance, Markowitz (1952) noted prevalent risk-proneness for lotteries involving losses among his subjects. This observation is also noted by

Kahneman and Tversky (1979), particularly their probabilistic insurance example. Figure 3.1 displays a typical von Neumann-Morgenstern utility function which is used to account for the joint risk averting/seeking behavior discussed above. The "convex" ("concave") regions correspond to risk-proneness (risk-aversion).

Fig. 3.1: A "typical" von Neumann-Morgenstern utility function



Since the von Neumann-Morgenstern utility function cannot be convex and concave at the same time, expected utility rules out concurrence of risk proneness and risk aversion within the same region of wealth levels. Whether this is actually the case is an empirical question that has yet to be fully investigated. There is however indirect evidence (MacCrimmon, et.al. 1972; Allais, 1977) to the contrary. Different theoretically equivalent procedures for the elicitation of von Neumann-Morgenstern utility, e.g., the certainty equivalent, the gain equivalent, and the chaining method, tend to yield different curves with opposing risk-propensities.

A difficulty with expected utility, one that touches on the largely unexplored area of problem representation and its effect on the decision maker's preference, is the controversy over the domain on which a utility function is defined. Should it be final wealth levels, the normatively compelling position as in Friedman and Savage (1948), Pratt (1964) and practically all the literature on the economics of uncertainty, or changes in asset position relative to some "customary" wealth level? Markowitz (1952) and others have observed that preferences are relatively independent of the current wealth levels. Another difficulty is related to the finding (Kahneman and Tversky, 1979) that preferences among two-stage lotteries may depend on whether the decision maker represents these lotteries in their simple equivalent forms.

This chapter expands on the issues introduced above without duplicating unduly the contents of other critiques already cited.

We shall explore the descriptive implications of our generalization of expected utility theory in the next chapter in light of the examples considered here.

3.2 SYSTEMATIC VIOLATION OF THE STRONG INDEPENDENCE PRINCIPLE

As a lead-in to a more general structure of lotteries, consider the four decision problems given in Figure 3.2.

| | |
|--|---|
| A_0 : \$1,000,000 for sure B_0 : 10/11 chance of \$5,000,000 1/11 chance of \$0 | A_H : 89/100 chance of \$5,000,000 11/100 chance of \$1,000,000 B_H : 99/100 chance of \$5,000,000 1/100 chance of \$0 |
| A_L : 11/100 chance of \$1,000,000 89/100 chance of \$0 B_L : 10/100 chance of \$5,000,000 90/100 chance of \$0 | A_I : \$1,000,000 for sure B_I : 10/100 chance of \$5,000,000 89/100 chance of \$1,000,000 1/100 chance of \$0 |

Figure 3.2: Four decision problems

Under the expected utility hypothesis, the only permissible patterns of choices are either A_H , A_I , A_L , A_0 or B_H , B_I , B_L , B_0 . If your choices are like most people's, you will have chosen A_H , A_I , B_L and A_0 , which is not consistent with the implications of expected utility. The choice of A_I and B_L constitutes the well known Allais

paradox. A lesser known paradox, the Allais ratio paradox, is given by the choice of B_L and A_0 . The violating choice of A_H and B_L has not been studied. The insight one gains from the structure in Figure 3.2, rather than simply considering separate binary lotteries, is that the violating pairs (A_I, B_L) , (B_L, A_0) and (A_H, B_L) are all derivatives of the basic violation, A_H, A_I, B_L, A_0 versus A_H, A_I, A_L, A_0 .

Several features of the structure in Figure 3.2 are worth noting. It is based on three consequences \$0, \$1,000,000 and \$5,000,000, denoted by L, I, and H respectively. The $A_x(B_x)$ alternative, where x stands for one of the consequences, L, I, H, is obtained from the $A_0(B_0)$ alternative by composition with consequence x at probability 89/100. This is illustrated in Figure 3.3 for the case $x = L$.

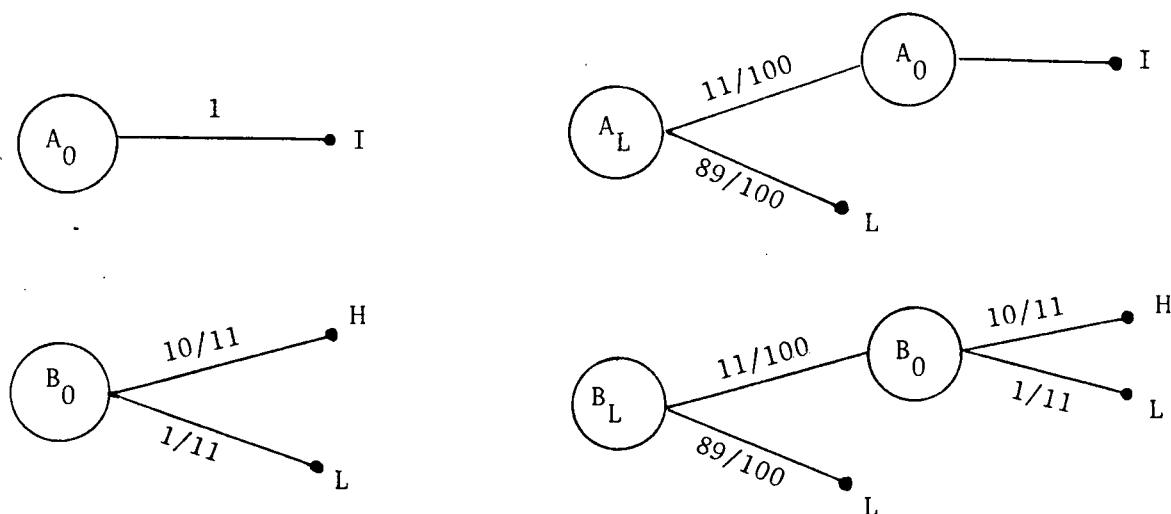


Figure 3.3: The composition of the A_L (B_L) lottery from A_0 (B_0)

Since $A_L(B_L)$ in Figure 3.3 has the same final outcomes and probabilities as $A_L(B_L)$ in Figure 3.2, these lotteries are equivalent.

Although for illustrative purposes, we have only considered the consequences \$0, \$1,000,000 and \$5,000,000 and the composition probability .89, it seems reasonable to expect violations of expected utility for other consequence values and other probability levels. This leads to a more general structure of decision problems, illustrated in Fig. 3.4. A_0 is a sure prospect of the intermediate consequence I. B_0 offers a q chance at the most preferred outcome H and a $1-q$ chance at the least preferred outcome L. The $A_x^\beta(B_x^\beta)$ alternative is obtained from the $A_0(B_0)$ alternatively by composing with the x -consequence at probability β .

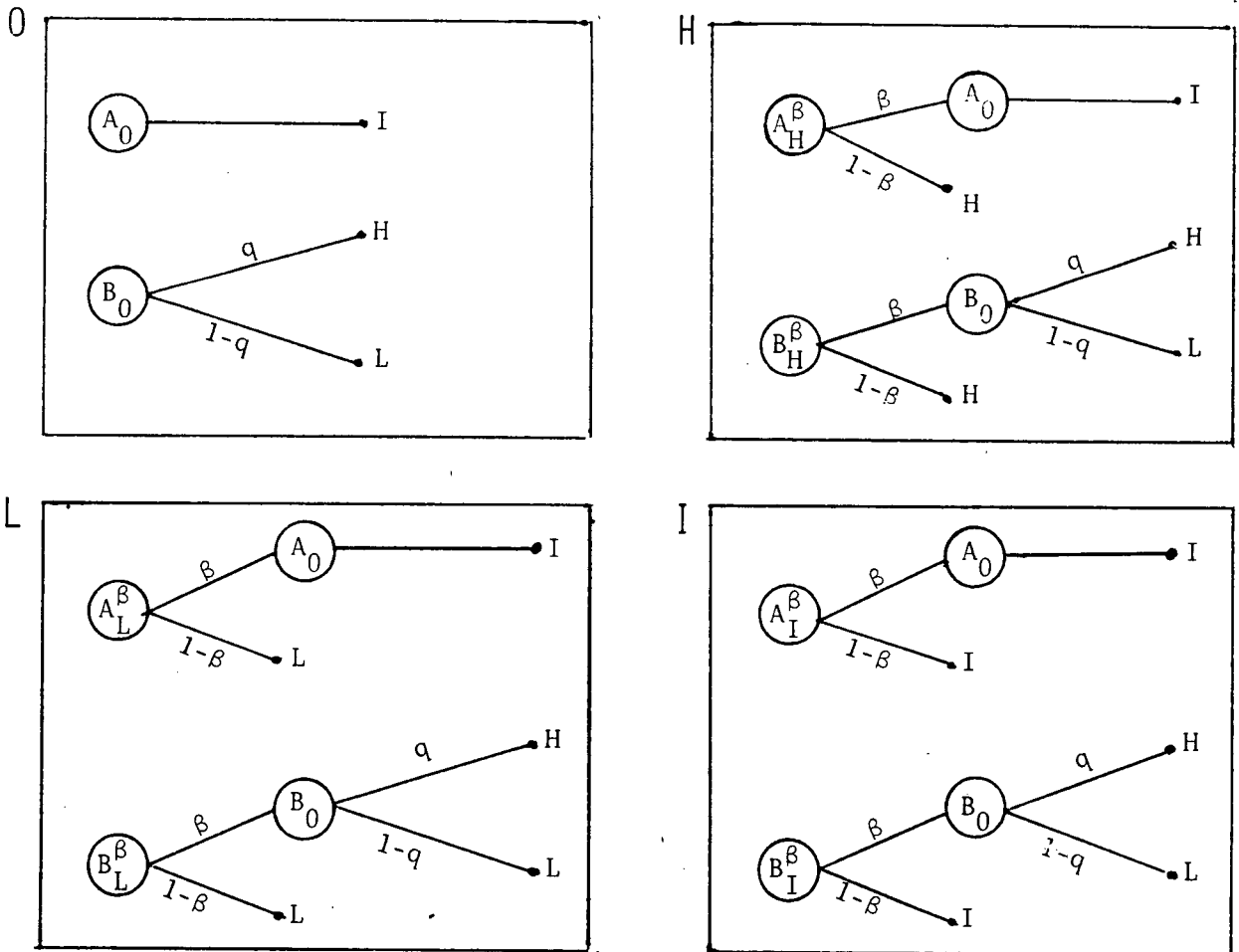


Figure 3.4: The HILO structure of three consequence lotteries

Note that the H (for "high"), I (for "intermediate") and L (for "low"), given in the boxes of Figure 3.4, exhaust the possible compositions from the basic problem (denoted as "0"). For ease of reference this will be called the "HILO" lottery structure.

Expected utility theory imposes some severe restrictions on the choices in this lottery structure. The strong independence principle requires that preference between two alternatives be preserved when each alternative is composed with a common alternative at the same probability. Since this is how the A_x^β and B_x^β alternatives are generated, it implies that the choice of alternative A_0 entails the choice of A_x^β while the choice of alternative B_0 entails the choice of B_x^β , for all values of x and β . Hence, the choice of an A alternative in one of the cases of the HILO structure and a choice of a B alternative in another case (such as the standard Allais choice of A_I, B_L) violates expected utility. In addition to these violations across problems, violations may occur within each problem since $A_x^{\beta_1}$ may be chosen for some particular level β_1 while $B_x^{\beta_2}$ may be chosen for some different level β_2 .

Although empirical results are not available for all the across-case and within-case combinations (since this lottery structure has not appeared in the literature), there are results for some of the particular cases. Table 3.1 summarizes the choices from the main empirical studies in terms of the HILO lottery structure. From this table it should be apparent that most of the effort has been devoted to studying various versions of the standard Allais paradox,

Table 3.1: Summary of Empirical Results Relevant to the HIL0 Lottery Structure

| Case: Violation Pattern | Empirical Studies | Parameters in the HIL0 Structure | | | | | Violation Percentage |
|---|-----------------------------|---------------------------------------|---|---|--|--|--|
| | | L | I | H | q | B | |
| I-L Case: A_L, B_L Standard Allais Paradox | Allais '79 | 0 Fr | 100 mill (old) Fr | 500 mill (old) Fr | 10/11 | 11/100 | 46% (n=101) |
| | MacCrimmon '65 | Bankruptcy Bankruptcy 0% return | 5% return very small return 35% return | 500% return high return 75% return | 10/11 qualitative description 10/11 | 11/100 11/100 11/100 | 39% (n=36) 42% (n=36) 36% (n=36) |
| | Moskowitz '74 | F grade | B+ grade | A grade | 10/11 | 11/100 | 32% (n=134) |
| | Stovic & Tversky '75 | 0% | $\$1 \times 10^6$ | $\$5 \times 10^6$ | 10/11 | 11/100 | 44% (n=78) |
| | MacCrimmon & Larsson '75 | \$0 | $\$1 \times 10^4$ $\$1 \times 10^5$ $\$1 \times 10^6$ | $\$5 \times 10^4$ $\$5 \times 10^5$ $\$5 \times 10^6$ | 10/11 10/11 10/11 | 11/100 11/100 11/100 | 21% (n=19) 21% (n=19) 32% (n=19) |
| | Kahneman & Tversky '79 | \$0 | \$2400 | \$2500 | 33/34 | 34/100 | 65% (n=72) |
| | Ashton '79 | 1%; bankruptcy 1% | 9% 8%; 18%; 28% | 45%; 450% 115%; 180% | 10/11 10/11 | 11/100 11/100 | 14% (n=83) 20% (n=49) |
| | Hagen '73 | 0 | 10000 Norw. Kr. | 20000 Norw. Kr. | .50 | .02 | 34% (n=52) |
| | MacCrimmon & Larsson '75 | \$0 | $\$1 \times 10^3$ $\$1 \times 10^6$ | $\$5 \times 10^3$ $\$5 \times 10^6$ | .80 .80 | .75; .50 .75; .50; .25; .10; .05 | 32%; 21% (n=19) 32%; 41%; 58%; 58%; 63% (n=19) |
| | Kahneman & Tversky '79 | \$0 No tour | \$3000 1 week tour | \$4000 3 week tour | .80 .50 | .25 .10 | 45% (n=95) 45% (n=72) |
| O-L Case: A_O, B_L | Hagen '79 | 0 | 5 mill Norw. Kr. | 25 mill Norw. Kr. | 10/11 | .99; .11 | 42% (n=53) |
| | MacCrimmon & Larsson '75 | \$0 | $\$1 \times 10^3$ $\$1 \times 10^6$ | $\$5 \times 10^3$ $\$5 \times 10^6$ | .80 .80 | .75; vs .50 .75; vs .50; .25; .10; .05 | 21% (n=19) 37%; 26%; 26%; 26%; 32% (n=19) |
| | Kahneman & Tversky '79 | \$0 | \$3000 | \$6000 | .50 | .90; .002 | 53% (n=66) |
| L-L Case: $A_L^B, B_L^{B'}$ ($B > B'$) | MacCrimmon & Larsson '75 | -\$5000 | -\$1000 | \$0 | .80 | .80; .20; .04 | 5%; 5%; 16% (n=19) |
| | Kahneman & Tversky '79 | -\$4000 | -\$3000 | \$0 | .80 | .25 | 50% (n=95) |
| O-H Case: $B_O^B, A_H^{B'}$ ($B > B'$) | Kahneman & Tversky '79 | -\$6000 | -\$3000 | \$0 | .50 | .90; .002 | 62% (n=66) |
| | MacCrimmon & Larsson '75 | -\$5000 | -\$1000 | \$0 | .80 | .80 vs .20; .04 .20 vs .04 | 11%; 21% (n=19) 21% (n=19) |
| H-H Case: $B_H^B, A_H^{B'}$ ($B > B'$) | MacCrimmon & Larsson '75 | -\$5000 | -\$1000 | \$0 | .80 | .80 vs .20; .04 .20 vs .04 | 11%; 21% (n=19) 21% (n=19) |

the I-L case.¹ Note that while the frequency of choosing the violating choices, A_I^β , B_L^β , varies across studies, the violation seems robust over quite different levels of consequences and probabilities. Receiving increasing attention recently has been the L-L pattern, $A_L^{\beta 1}$, $B_L^{\beta 2}$ (including the special case 0-L, of choices A_0 , $B_L^{\beta 2}$).

The only studies which have considered several cases simultaneously are those of MacCrimmon and Larsson (1975) and Kahneman and Tversky (1979). The former study introduced the 0-H and the H-H cases in the context of negative outcomes and is the only attempt to map patterns of preferences between A_L^β and B_L^β for various levels of β (including the special "0" case of $\beta = 1.0$) and various levels of the intermediate outcome, I.

It seems clear from Table 3.1 that our understanding of actual choices for the decision problems in Figure 3.4 is incomplete. Studies thus far conducted have covered the I-L, L-0, and L-L cases, for gains, and the H-0 and H-H cases, for losses. With the HILO lottery structure, there are potentially 6 distinct binary violation patterns, "I-L", "H-I", "I-0", "L-0", "H-0" and "H-L", across problems and three binary violation patterns, "H-H", "I-I" and "L-L" within problems. The cases "H-I", "I-0", "I-I" and "H-L" remain unexplored. To obtain a more complete picture, a systematic study of choices related to Figure 3.4 is needed.

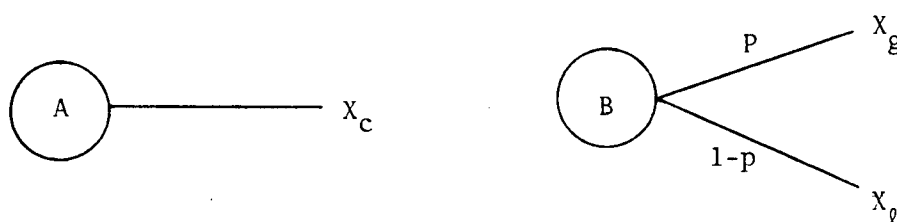
¹ The terminology "I-L case" refers to the pair of lotteries comprising the I case being presented in conjunction with the lotteries of the L case.

3.3 CONCURRENCE OF RISK SEEKING AND RISK AVERTING BEHAVIOR

We shall not discuss here the existence of non-overlapping risk-seeking and risk-averting regions of a utility function (see e.g. Fig.3.1) corresponding to the purchase of lottery tickets, purchase of insurance and greater risk-seeking propensity for lotteries involving losses. This has been given adequate coverage elsewhere (Friedman and Savage, 1948; Markowitz, 1952; Kahneman and Tversky, 1979). Instead, we focus our attention on the possible concurrence of risk proneness and risk aversion within the same region; thus negating any explanation based on modifications of the von Neumann-Morgenstern utility function.

Several measurement procedures to elicit a decision maker's von Neumann-Morgenstern utility function are based on lottery comparison of the sort given in Figure 3.5.

Figure 3.5: Standard lottery comparison



Lottery A is a sure consequence of X_c , an amount greater than X_l but less than X_g . Lottery B is a p chance of getting X_g and a $1-p$ chance of getting X_l . If we fix the amounts in lottery B at X_l^0 and X_g^0

respectively but allow X_c and p to vary, then the pairs of numbers (X_c, p) such that lottery A is indifferent to lottery B define a function which we denote by $\bar{u}_c(X)$, i.e. lottery A is indifferent to lottery B whenever $p = \bar{u}_c(X_c)$. We denote an affine transformation of $\bar{u}_c(X)$ by $u_c(X)$, i.e. $u_c(X) = a\bar{u}_c(X) + b$, for some a, b with $a > 0$. Suppose the decision maker is an expected utility maximizer with von Neumann-Morgenstern utility function $u(X)$. Then

A indifferent to B implies that

$$\begin{aligned} u(X_c) &= \bar{u}_c(X_c)u(X_g^0) + (1-\bar{u}_c(X_c))u(X_\ell^0) , \text{ or alternatively} \\ u(X_c) &= [u(X_g^0) - u(X_\ell^0)]\bar{u}_c(X_c) + u(X_\ell^0) , \end{aligned} \quad (3.1)$$

which is an affine transformation of $\bar{u}_c(X_c)$.

Therefore, the function $u_c(X)$ is a von Neumann-Morgenstern utility function. The above measurement procedure is usually known as the *Certainty Equivalent* method.

We can alternatively fix the sure amount in A at X_c^0 and the loss amount in B at X_ℓ^0 and obtain the pairs $(X_g, p_g(X_g))$ such that A remains indifferent to B. Applying expected utility again, we obtain the following relation.

$$\begin{aligned} u(X_c^0) &= p_g(X_g)u(X_g) + (1-p_g(X_g))u(X_\ell^0) , \text{ or alternatively} \\ u(X_g) &= \frac{u(X_c^0) - u(X_\ell^0)}{p_g(X_g)} + u(X_\ell^0) . \end{aligned} \quad (3.2)$$

Since $u(X)$ is related to $p_g^{-1}(x)$ through an affine transformation, $p_g^{-1}(X)$ is also a von Neumann-Morgenstern utility function; and this

elicitation procedure is called the *Gain Equivalent* method.

Similarly, we can determine the von Neumann-Morgenstern utility using the *Loss Equivalent* method by fixing X_c^0 and X_g^0 to obtain the pairs $(X_\ell, p_\ell(X_\ell))$ such that a $p_\ell(X_\ell)$ chance of X_g versus getting X_ℓ otherwise is indifferent to getting X_c^0 for sure. We apply expected utility again and obtain:

$$u(X_\ell) = u(X_g^0) - \frac{u(X_g^0) - u(X_c^0)}{1 - p_\ell(X_\ell)} \quad (3.3)$$

Therefore, $\frac{-1}{1 - p_\ell(X)}$ is a von Neumann-Morgenstern utility function.

The certainty, gain and loss equivalent methods are obtained from Fig. 3.5 by holding constant the pairs (X_ℓ, X_g) , (X_ℓ, X_c) and (X_c, X_g) respectively. If instead we hold p^0 constant, then there are three remaining candidates for additional methods known collectively as the *Chaining Methods* which are obtained by holding constant the pairs (p^0, X_ℓ^0) , (p^0, X_c^0) and (p^0, X_g^0) respectively. Of the three cases, we describe only the first i.e. fixing (p^0, X_ℓ^0) , which is more often used in practice. Beginning with an amount X_1 larger than X_ℓ^0 , determine an amount X_2 such that a sure consequence of X_2 in A would be indifferent to p^0 chance of getting X_1 and obtaining X_ℓ^0 otherwise in B. Determine a third value X_3 by replacing X_1 by X_2 and repeating the process. Thus, we obtain a decreasing sequence (X_1, X_2, X_3, \dots) with the following property,

$$u(X_{i+1}) = p^0 u(X_i) + (1 - p^0) u(X_\ell^0). \quad (3.4)$$

Assigning arbitrary values to $u(X_\ell^0)$ and a greater value to $u(X_1)$,

we observe that expression (3.4) determines the von Neumann-Morgenstern utility on the decreasing set of points (X_1, X_2, X_3, \dots) . We shall denote the utility function obtained using the chaining method with fixed probability p^0 and fixed loss amount X_ℓ^0 by $u_{p^0}(X)$.

As we have noted earlier, if the decision maker is a 'true' expected utility maximizer, then the utility functions u_c , u_g , u_ℓ and u_{p^0} would be affine transformations of each other so that any one of them suffices. Allais (1979, Appendix C), in an experiment conducted in 1952 found that the u_c and $u_{\frac{1}{2}}$ curves obtained from the same subjects are generally very different. A fairly typical plot is given in Figure 3.6. Note that under the expected utility hypothesis, the convex region near $X=0$ of u_c is not compatible with the global concavity of $u_{\frac{1}{2}}$. Behaviorally, the risk seeking region of u_c corresponds to our intuition about the psychology of lottery purchase -- people tend to forgo a small certain amount in favour of a small chance of a large gain; while the concavity of $u_{\frac{1}{2}}$ reaffirms the reluctance of individuals to engage in symmetric bets.

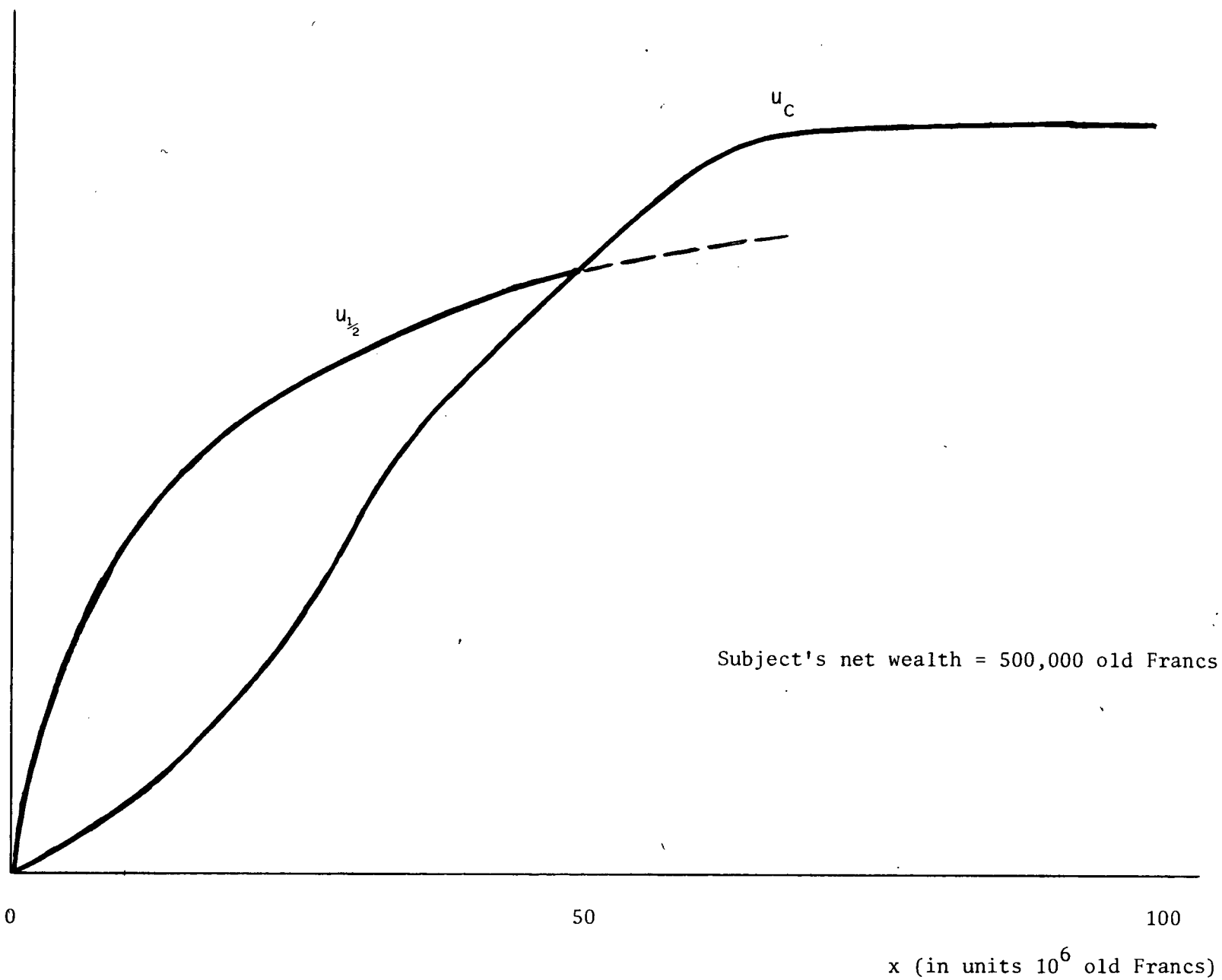


Figure 3.6: Examples of u_c and $u_{1/2}$ (based on data from Allais (1979), Appendix C)

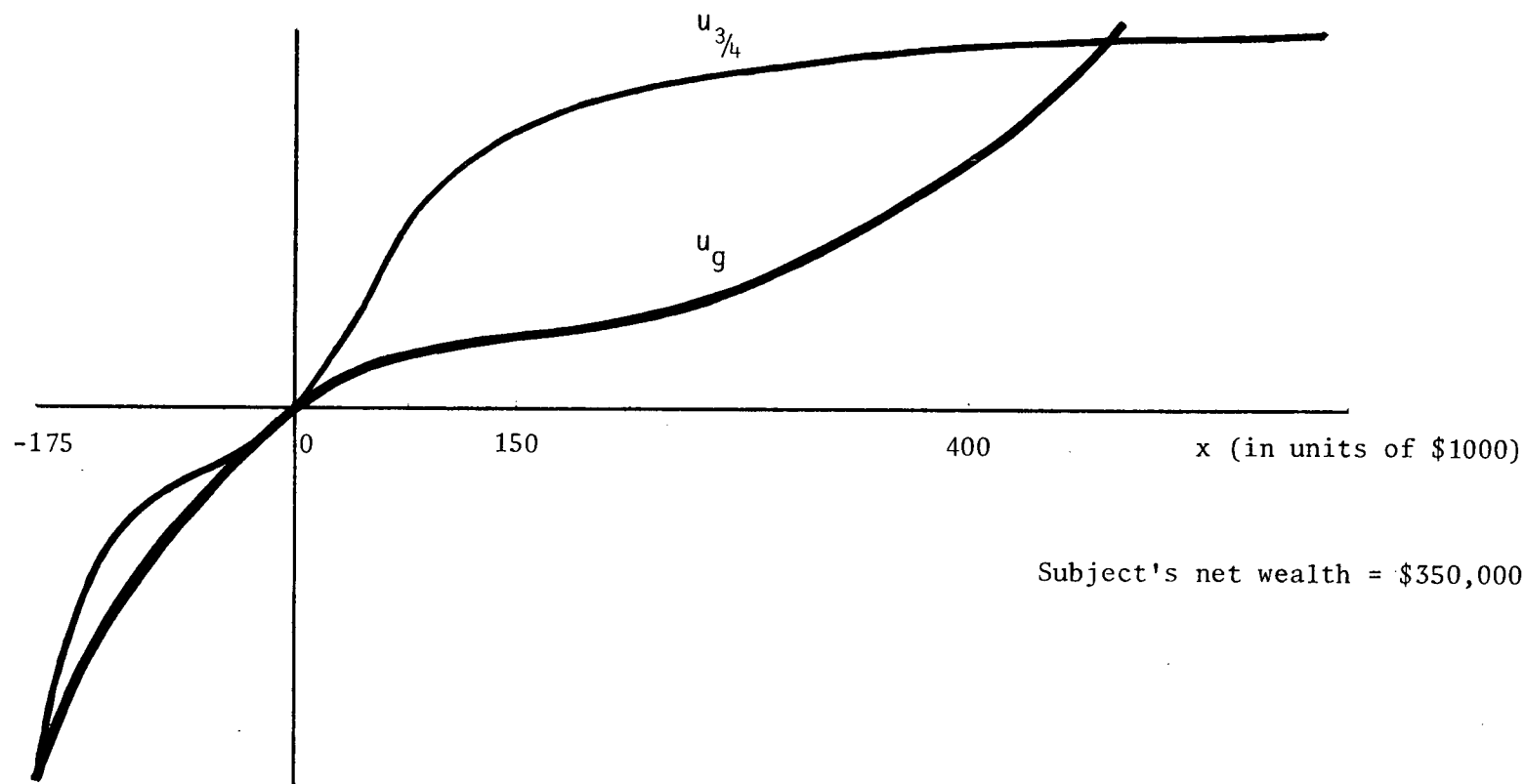


Figure 3.7: Examples of $u_{3/4}$ and u_g (based on data from MacCrimmon et. al. (1972))

In an ongoing study on the risk attitudes of top-level business managers carried out by MacCrimmon and others (1972), chaining and gain equivalents were among the methods used to assess von Neumann-Morgenstern utility functions. Fig. 3.7 displays a typical pair of curves $u_{3/4}$ and u_g obtained from a subject using the chaining method and gain equivalent method respectively. Note again that the convexity of $u_{3/4}$ near $X=0$ is inconsistent, under the expected utility hypothesis, with the concavity of u_g at the same region. The same inconsistency also applies, though in the opposite direction, to the convex region of u_g versus the concavity of $u_{3/4}$ beyond its initial convex region. Even though the empirical evidence on the concurrence of risk proneness and risk aversion within the same range of wealth levels is scant and fragmentary, what we already know about the actual application of different methods to elicit utility functions suggests that expected utility does not account for the results. This suggests that further investigation of the concurrence of risk proneness and risk aversion and especially mutual incompatibility of the different measurement procedures to obtain von Neumann-Morgenstern utility functions is warranted.

3.4 SOME PROBLEMS WITH PROBLEM REPRESENTATION

Problem representation and its influence on preferences is a relatively untouched area of research on decision-making. Normatively, a decision maker's preference should not depend on the way alternatives are perceived or represented as long as it does not affect the desirability of the underlying consequences of his alternatives. That this may not be the case is demonstrated by Kahneman and Tversky (1979) through a class of phenomena termed Isolation Effects. Consider the choice between A and B in Figure 3.8.

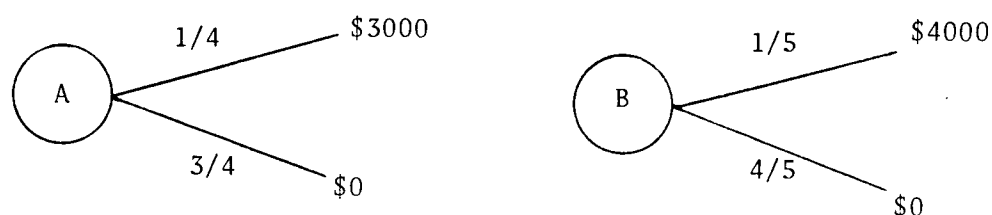


Fig. 3.8: Graphical representation of two lotteries

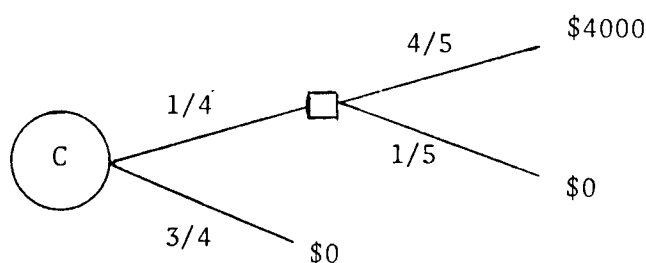


Fig. 3.9: A sequential representation of lottery B

If you are an expected utility decision maker, then your preference does not depend on how the probabilities of final outcomes are obtained, so that lottery C in Fig. 3.9 is equivalent to lottery B in Fig. 3.8, i.e. preference between A and B should be in the same direction as preference between A and C. Kahneman and Tversky found for one group of subjects ($n = 95$), 65% prefer lottery B to lottery A. However, the modal preference pattern between lottery A and lottery C for another group of subjects is found to be the opposite (78% prefer C to A; with $n=141$). The problem description that elicited the modal preference for lottery C versus lottery A is reproduced below. (4000,.80) refers to a lottery that pays \$4000 with .8 probability and \$0 with .2 probability, and (3000) denotes the sure consequence of \$3000.

Consider the following two-stage game.

In the first stage, there is a probability of .75 to end the game without winning anything, and a probability of .25 to move into the second stage. If you reach the second stage you have a choice between

(4000,.80) and (3000)

Your choice must be made before the game starts, i.e., before the outcome of the first stage is known.

The problem description above focuses the subjects' attention on the choice between (4000,.80) and (3000) rather than the common outcome \$0 with the same probability .75. Kahneman and Tversky conjectured that most subjects then ignore the common outcome - probability component under the above problem representation so that their choice becomes identical to that between (4000,.80) and (3000).

Another kind of isolation effect considered by Kahneman and Tversky is related to Markowitz (1952)'s observation that preferences are relatively independent of current wealth levels. They presented the following problems to two different groups of subjects.

Problem 1. In addition to whatever you own, you have been given 1000. You are now asked to choose between

A: (1000,.50) and B: (500)

Problem 2. In addition to whatever you own, you have been given 2000. You are now asked to choose between

C: (-1000,.50) and D: (-500)

The majority of subjects chose A in Problem 1 and B in Problem 2. Note, however, that in terms of final outcomes, the two choice problems are equivalent i.e. you are either \$1500 richer if you choose B or D, or you have even chance of ending up with \$1000 or \$2000 more if you choose A or C. People, however, seem to perceive Problem 1 as a choice between (1000,.50) and (500), and Problem 2 as a choice between (-1000,.50) and (-500), with the lump sums of \$1000 in Problem 1 and \$2000 in Problem 2 safely integrated into their current wealth levels.

As an alternative to the final outcome position normally associated with expected utility, Kahneman and Tversky proposed that people perceive outcomes as gains and losses relative to some neutral reference point. Allowing the reference point to be determined by the decision maker in the context of the choice situation he faces, expected utility, with a utility function defined on changes in asset position relative to the reference point, is compatible with the modal preferences in

Problems 1 and 2. The reference point need not be the status quo especially when the choice situation involves a sure gain of \$1000.

3.5 SUMMARY

We have reviewed briefly some of the literature on empirical evidence that contradicts the implications of expected utility. It has been classified under the headings: systematic violations of the strong independence principle (section 3.2), concurrence of risk proneness and risk aversion (section 3.3), and some problems with problem representation (section 3.4). The phenomena considered include the Allais paradox and its various modifications, incompatibility among different methods of measuring a von Neumann-Morgenstern utility function, and Kahneman and Tversky's isolation effects.

In the next chapter, we generalize expected utility theory by applying the representation theorems of Part I as a theory of choice. We then explore the descriptive implications of our generalization with respect to the phenomena considered in this critique.

A NEW THEORY

We develop in this chapter a new theory of choice called alpha utility theory which generalizes expected utility by interpreting, in sections 1 and 2, the representation theorems of Chapters 1 and 2 in terms of choice among lotteries. We explore in section 3 some normative implications of our theory including consistency conditions with stochastic dominance and local and global risk aversion. The question of descriptive validity is considered in section 4 where we show that alpha utility is compatible with the phenomena reviewed in the critique of expected utility (Chapter 3). We end the chapter with a comparison of our theory with two other alternative theories of choice.

4.1 INTERPRETING MEAN VALUE AS CERTAINTY EQUIVALENT

In Chapter 1, we proved a representation theorem of a mean value functional, M , for probability distributions subject to a necessary and sufficient set of axioms. The present section explores the implications of our representation theorem for choice among lotteries by interpreting mean values as certainty equivalents.² In the ensuing discussion, we assume that lotteries, defined on some single-attribute consequence set, e.g., monetary gains, can be represented by probability distributions defined on the real line. (The case of more general consequence space is considered in section 4.2)

The decision maker is assumed to have complete and transitive preference over lotteries. The following axiom asserts that he is able to assign, corresponding to any lottery F , a certainty equivalent

² This was discussed in Chew (1979).

$M(F)$ which is an amount such that the decision maker is indifferent between getting it for sure and taking the lottery F . D denotes the space of probability distributions (lotteries) defined on the real line.

Axiom M0: Existence $\forall F \in D$, $M(F)$ exists.

Axiom M0 rules out infinite certainty equivalents thus pre-empting any possibility of a St. Petersburg type paradox.

The following five axioms M1-M5 are taken directly from Chapter 1. We did not state existence (M0) as an axiom in Chapter 1 because it is not an intrinsic property of mean values. For example, the arithmetic mean does not always exist.

Axiom M1: Certainty Consistency $M(\delta_x) = x \quad \forall x \in R$.

It is difficult to take issue with Axiom M1 which requires that the certainty equivalent of δ_x i.e. getting x for sure, is x . Another normatively appealing axiom is the following:

Axiom M2: Betweenness $\forall F, G \in D$, if $M(F) < M(G)$, then
 $\forall \beta \in (0,1)$, $M(\beta F + (1-\beta)G) \in (M(F), M(G))$.

Axiom M2 requires that the certainty equivalent of a mixture of two lotteries be intermediate in value between the certainty equivalents of the respective lotteries. The next axiom weakens the axiom of quasilinearity of Hardy, Littlewood and Polya or the "substitution of lotteries" principle of Pratt, Raiffa and Schlaifer.

Axiom M3: Weak Substitution $\forall F, G \in D$, if $M(F) = M(G)$,
 then $\forall \beta \in (0,1) \exists \gamma \in (0,1)$
 $\ni \forall H \in D, M(\beta F + (1-\beta)H) = M(\gamma G + (1-\gamma)H)$.

Suppose F and G are two lotteries with the same certainty equivalents i.e. indifferent to each other. Consider the mixtures or compound lotteries $\beta F + (1-\beta)H$ and $\gamma G + (1-\gamma)H$ of the respective lotteries, with a third lottery H at probabilities β and γ . Weak substitution weakens the substitution principle in the sense that the mixture probabilities β and γ that preserve equality of certainty equivalents or indifference need not be the same. However, it requires that these mixture probabilities once determined be independent of the third lottery H .

The next two axioms are technical assumptions which are introduced to ensure that our notion of certainty equivalent is well behaved relative to certain limiting operations on probability distributions. Axiom M4 permits the approximation of lotteries with numerous discrete outcomes by a continuous distribution e.g. uniform distribution.

Axiom M4: Continuity If $\{F_n\}_{n=1}^{\infty}$ converges in distribution to F
 and F has compact support, then
 $M(F) = \lim_{n \rightarrow \infty} M(F_n)$.

For a more technical discussion of Axiom M4 see Chapter 1. Our last axiom stated below asserts that the certainty equivalent of a lottery F_K , truncated by the interval K , is 'close' to that of the original lottery F if the interval of truncation K is sufficiently large.

Axiom M5: Extension Let $\{K_n\}_{n=1}^{\infty}$ be an increasing family of compact intervals such that $\lim_{n \rightarrow \infty} K_n = R$.
Then $M(F) = \lim_{n \rightarrow \infty} M(F_{K_n})$, $\forall F \in D$.

We are now ready to restate Theorem 1.3 of Chapter 1 in terms of certainty equivalents.

Suppose a decision maker assigns a certainty equivalent $M(F)$ corresponding to a lottery F , then the assignment of $M(F)$ satisfies Axioms M1-M5 if and only if there exist a strictly positively valued function α and a strictly increasing function v such that $\forall F \in D$,

$$M(F) = v^{-1} \left(\int_R \alpha v dF / \int_R \alpha dF \right) . \quad (4.1)$$

Moreover, if (α^*, v^*) is another pair of functions that satisfies condition (4.1), then, for every interval $[A, B] \subset R$, there exist constants a, b, c, k with $a, c, k > 0$ such that $\forall x \in [A, B]$,

$$v^*(x) = a \frac{k[v(x) - v(A)]}{k[v(x) - v(A)] + v(B) - v(x)} + b ,$$

$$\text{and } \alpha^*(x) = c\alpha(x) [k(v(x) - v(A))] + v(B) - v(x) . \quad (4.2)$$

Note that the expression for certainty equivalent (4.1) is more general than the corresponding expression for expected utility,

$$M(F) = v^{-1}(\int_R v dF) \quad (4.3)$$

which results when α is constant. The uniqueness relation (4.2) subsumes as special cases, an affine transformation for v and a positive scalar transformation for α . To avoid the possibility of any St. Petersburg type paradox, we have asserted through Axiom M0 that the certainty equivalent $M(F)$ must be finite for any lottery F . We can then apply Corollary 1.3 in Chapter 1 to conclude that either v is bounded or $\alpha \cdot v$ is bounded. Thus, we have obtained a generalization of expected utility, called alpha utility, in the sense that the decision maker's preference over lotteries is represented by a more general expression for his certainty equivalent characterized by an additional alpha-function. We shall defer discussion of the actual shapes of α and v functions till the sections after next. The next section contains a parallel development of the above theory using the representation theorems of Chapter 2.

4.2 REPRESENTATION OF A PREFERENCE BINARY RELATION

In the preceding section, we developed alpha utility theory for lotteries defined on a single-attribute consequence set, e.g., monetary lotteries, via an axiomatization of the decision maker's assignment of certainty equivalents. We showed, using the mean value representation theorems of Chapter 1, that the certainty equivalent $M(F)$ corresponding to a lottery F is given by:

$$M(F) = v^{-1}(\int_R \alpha v dF / \int_R \alpha dF) \quad (4.4)$$

for some strictly positive function α and strictly increasing function v . When a lottery F is preferred to a lottery G , the certainty equivalent $M(F)$ is greater than that of G , $M(G)$. It follows (since v^{-1} is order-preserving) that

$$\int_R \alpha v dF / \int_R \alpha dF > \int_R \alpha v dG / \int_R \alpha dG. \quad (4.5)$$

Observe that the preference represented by the above functional is more general than that of expected utility, $\int_R v dF$, which is a special case of (4.5) when α is constant.

The present section applies the representation theorems of Chapter 2 to provide an alternative axiomatization of alpha utility theory.³ Unlike the certainty equivalent approach, we do not limit ourselves to choice situations where the range of consequences are monetary values or some quantity of certain commodity. For example, the relevant consequence set for a child contemplating whether to steal a cake from a bakery may be: status quo, having the cake, getting caught in the process. We denote by

³ This was discussed in Chew and MacCrimmon (1979a).

X ($=\{x,y,z,\dots\}$) the consequence set corresponding to a choice situation and L_X ($=\{P, Q, R,\dots\}$) the space of simple probability measures defined on X (cf. Chapter 1). A simple probability measure is completely specified by knowledge of the probabilities of occurrence of a finite number of consequences. The child in the above example may feel that he has an even chance of getting the cake without being caught. Simple probability measures are convenient representations of actual lotteries or risky decisions when the probabilities of occurrence of the underlying consequences can be subjectively estimated or determined based on symmetry considerations, e.g., a game of craps or roulette. We denote the alternative of obtaining some consequence x for sure by δ_x . A finite lottery P is then represented as a probability weighted combination of sure consequences:

$$P \equiv \sum_{i=1}^n p_i \delta_{x_i} \quad (4.6)$$

where p_i is the probability of occurrence of the consequence x_i .

Our representation functional for simple probability measures corresponding to that obtained from the certainty equivalent approach (see expression (4.5)) is:

$$E(\alpha v, P) / E(\alpha, P), \quad (4.7)$$

where α and v are functions on X , and $E(\cdot, P)$ denotes taking expectation of a function with respect to the simple probability measure P .

The theorems of Chapter 2 show that the axioms on the strict prefer-

ence binary relation ' $<$ ' of a decision maker stated below are necessary and sufficient for representing his preference via expression (4.7).

Axiom U1: Ordering $<$ is a weak order.

The strict preference relation $<$ is a weak order if it is asymmetric (i.e., if a lottery P is strictly preferred to a lottery Q , then the converse does not hold), and negatively transitive (i.e., if a lottery P is not strictly preferred to a lottery Q which is in turn not strictly preferred to another lottery R , then P is not strictly preferred to the lottery R). Both asymmetry and negative transitivity seem like basic consistency requirements for a strict preference relation that few would want to violate.

Axiom U2: Solvability $\forall P, Q, R \in L_X, P < Q \text{ and } Q < R$
 $\Rightarrow \exists \beta \in (0,1)$
 $\ni \beta P + (1-\beta)R \sim Q.$

Axiom U2 says that whenever a lottery Q is between two lotteries P and R in preference, then there is a mixture between P and R which is indifferent to Q . The reasonableness of the above axiom stems from the intuition that a mixture between two lotteries is always intermediate in preference between them (see the Betweenness axiom in section 4.1 and also Axioms 3:B:a and 3:B:b in von Neumann and Morgenstern (1947)). It follows that a mixture with a greater probability weight on the better lottery should be preferred to another mixture with a smaller probability weight on the better lottery. This is the substance of the following axiom.

Axiom U3: Monotonicity $\forall P, Q \in L_X, P \prec Q$
 $\Rightarrow \beta P + (1-\beta)Q \prec \gamma P + (1-\gamma)Q \text{ for } 0 \leq \gamma < \beta \leq 1.$

The axioms discussed thus far are standard normative properties common to expected utility theory. The next axiom which weakens the substitution principle or its close counterpart, the strong independence principle, is the only departure.

Axiom U4: Weak Independence $\forall P, Q \in L_X, P \sim Q$
 $\Rightarrow \forall \beta \in (0,1), \exists \gamma \in (0,1) \ni \forall R \in L_X,$
 $\beta P + (1-\beta)R \sim \gamma Q + (1-\gamma)R.$

Weak Independence is a restatement of the Weak Substitution axiom (M3) for certainty equivalents. Given two lotteries that are indifferent to each other, Weak Independence allows for different probabilities in composing each of these lotteries with a third lottery to preserve indifference. However, these mixture-probabilities once determined must be independent of the third lottery. We are now ready to interpret the representation theorems of Chapter 2.

A decision maker, whose preference among finite lotteries defined on a consequence set X satisfies Axioms U1-U4, chooses as if maximizing the functional

$$E(\alpha v, \cdot) / E(\alpha, \cdot)$$

for some strictly positive-valued function α and real valued function v both defined on the consequence set X .

When the α function is constant, the representation $E(\alpha v, \cdot)/E(\alpha, \cdot)$ becomes the expected utility representation $E(v, \cdot)$ with v assuming the role of a von Neumann-Morgenstern utility function.

Although the approach taken in this section does not require the consequence set to be single-dimensional, the certainty equivalent approach has the advantage of being able to treat general probability distributions, e.g., continuous random variable or even unbounded ones like the normal distribution. In the next two sections where we discuss the normative and descriptive implications of alpha utility theory, we shall represent lotteries in terms of either probability distributions or simple probability measures. Representing lotteries with probability distributions is more appropriate when we deal with numerical outcomes, especially if the outcome can take on continuous values. Otherwise, simple probability measures would be used.

4.3 NORMATIVE IMPLICATIONS

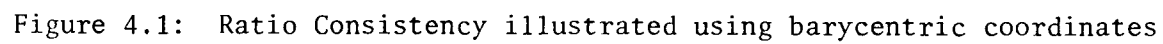
Despite growing evidence demonstrating its descriptive inadequacy, expected utility remains dominant mainly because of the normative appeal of its underlying postulates and the elegance with which expected utility characterizes stochastic dominance (increasing utility function), local (Arrow-Pratt index) and global (concave utility function) risk aversion. We show in this section that alpha utility is also an attractive

normative theory before considering its descriptive relevance in the following section.

4.3.1 Ratio Consistency

As we noted earlier, except for the Weak Substitution axiom (M3) or equivalently the Weak Independence axiom (U4), the other axioms of alpha utility are standard properties common to expected utility. We showed in Chapter 1 (Lemma 1.2) and Chapter 2 (Lemma 2.2) that Weak Independence together with Monotonicity implies the Ratio Consistency Property (Chapter 1, Property 5; Chapter 2, Definition 2.6) which is crucial to our proofs and the assessment of the α function. The Ratio Consistency property has been given a rather striking geometrical interpretation by Weber (1980).⁴ Consider a simplex formed by the three lotteries P, Q and R in Figure 4.1. Each lottery X in the simplex is specified by its barycentric (areal) coordinates $(\beta_1, \beta_2, \beta_3)$ given by the areas of the triangles XQR, RPX and XPQ. Assume that the area of PQR is 1, X then represents a probability mixture $\beta_1 P + \beta_2 Q + \beta_3 R$ among the three vertices P, Q and R. The lottery Y(Z) is a mixture between P(Q) and R with $\beta(\gamma)$ weight on P(Q) and $1 - \beta(1 - \gamma)$ weight on R.

⁴ Ratio Consistency was originally Axiom U5 (see Chew and MacCrimmon, 1979a). It was also implicit in a stronger statement of Axiom M3 (see Chew, 1979). Weber (1980) demonstrated via "an approach through the analysis of iso-preference sets" the redundancy of Ratio Consistency with respect to the other axioms by showing that Theorem 2.1 holds without assuming Ratio Consistency. This motivated the statements of Lemma 1.2 and Lemma 2.2.



Mixing Y and Z with weights varying from 0 to 1 generates the line segment YZ. Suppose the lotteries P and Q are indifferent to each other, then Monotonicity implies that the line segment PQ is an isopreference set. For each Y on PR, Weak Independence implies there is an Z on QR such that Y and Z are indifferent. Hence YZ is an isopreference. Project YZ to meet PQ extended at the point O with coordinates $(s, 1-s, 0)$ (Note that either $s > 1$ (O is on the left of P) or $s < 0$ (O is on the right of Q)). The Ratio Consistency property simply requires all other isopreference sets such as Y'Z' joining points on PR to points on QR to originate at the same point O. That this is the case follows from the observation that

$$\Delta OZR = \Delta OYR + \Delta YZR, \quad (4.8)$$

which is given by

$$\gamma s = \beta(s-1) + \gamma\beta. \quad (4.9)$$

Solving for γ gives

$$\gamma = \beta(s-1)/(s-\beta). \quad (4.10)$$

Let $\tau = s-1/s$. Expression (4.10) becomes

$$\frac{\gamma/1-\gamma}{\beta/1-\beta} = \tau. \quad (4.11)$$

A similar observation for the line segment Y'Z' gives

$$\frac{\gamma'/1-\gamma'}{\beta'/1-\beta'} = \tau = \frac{\gamma/1-\gamma}{\beta/1-\beta}. \quad (4.12)$$

Hence, line segments originating from the same point O satisfy the Ratio Consistency property (4.12). For expected utility, the mixture probabilities β and γ such that Y is indifferent to Z are always equal; so that the constant τ equals unity. Geometrically, the line segments YZ and $Y'Z'$ are parallel.

The simplicial representation can now be used as in Weber (1980) to provide an alternative proof of Lemma 1.2 or equivalently Lemma 2.2 without having to solve the functional equation (1.6). The basic idea is to show, once an isopreference say YZ is found, that all line segments such as $Y'Z'$ originating from the point O of the intersection between the extensions YZ and PQ are also isopreferences. Monotonicity which implies that these isopreferences are non-intersecting then ensures that there are no other isopreferences. Lemma 2.2 is proved below using the simplicial approach.

Lemma 2.2: Ordering (U1), Monotonicity (U3) and Weak Independence (U4) implies Ratio Consistency (Definition 2.6).

Proof: Suppose $P \sim Q$ but not $(P \sim R)$ in Figure 4.2.

Then from Monotonicity, PQ is an isopreference.

Corresponding to a mixture $Y \equiv \beta P + (1-\beta)R$, Weak Independence implies that there is a mixture $Z \equiv \gamma Q + (1-\gamma)R$ such that $Y \sim Z$, so that YZ is also isopreference. Let O be the point of intersection of the extensions of YZ and PQ . (Suppose without loss of generality that O is on the left of P).

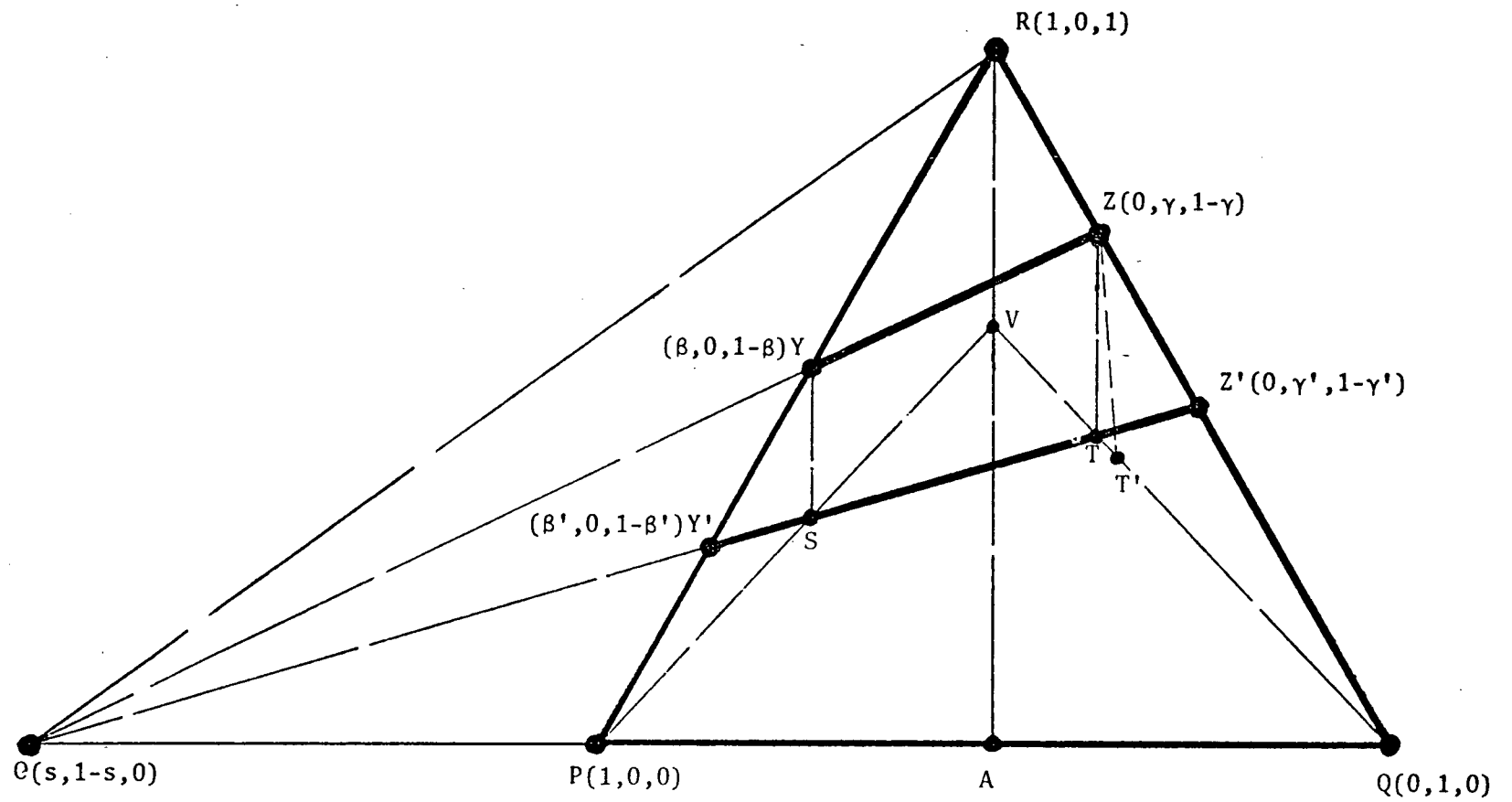


Figure 4.2: Geometric proof of the Ratio Consistency property

Consider another line segment $Y'Z'$ below YZ originating from the same point O . We shall show that $Y'Z'$ is also an isopreference and by the preceding discussion satisfies the Ratio Consistency property (4.12). Draw the median from R to bisect PQ at A . Draw a line from Y parallel to RA and intersecting $Y'Z'$ at S . Produce PS to meet RA at V . Complete the triangle PVQ by connecting V and Q . Denote by T the intersection between VQ and $Y'Z'$. Draw a line from Z parallel to RA and intersecting VQ at T' . To see that OST' is collinear and hence T' must be the same as T , view the figure in three dimensions with V the top vertex of a tetrahedron with base PQR . The plane through the parallel lines YS and ZT' contains O , S and T' . The plane through PV and Q also contains O , S and T' . The conclusion that $Y'Z'$ is also an isopreference follows from applying Weak Independence to $S \equiv \beta P + (1-\beta)V$ and $T \equiv \gamma Q + (1-\gamma)V$.

A similar argument establishes the result for the case $Y'Z'$ above YZ .

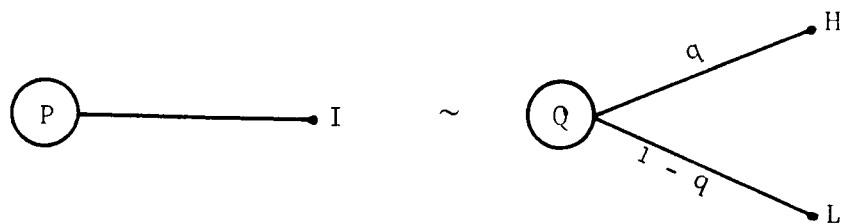
Q.E.D.

4.3.2 Assessment

The derivation of a procedure for eliciting alpha utility functions directly from the proofs of the representation theorems is a feature that is shared with expected utility and sets alpha utility apart from other alternative theories in the literature. This is illustrated below via a simple consequence set X with distinct outcomes L, I , and H arranged in ascending order of preference. Chapter 3 section 1 contains a discussion of the systematic violations of the Strong Independence principle by stating the empirical studies on the Allais paradox in terms of lotteries defined on such a 3-outcome consequence set. Suppose an alpha utility decision maker chooses among lotteries defined on such a consequence set, then we can measure his α and v functions in the following way:

Set the v -values for the worst and the best consequences to 0 and 1 respectively, i.e., $v(L) = 0$ and $v(H) = 1$. The v -value of the intermediate outcome I is given by the probability q such that the decision maker is indifferent between the sure consequence I and a q chance of obtaining H and $1-q$ chance of obtaining L , as illustrated in Fig. 4.3.

Figure 4.3: Probability Equivalent Method

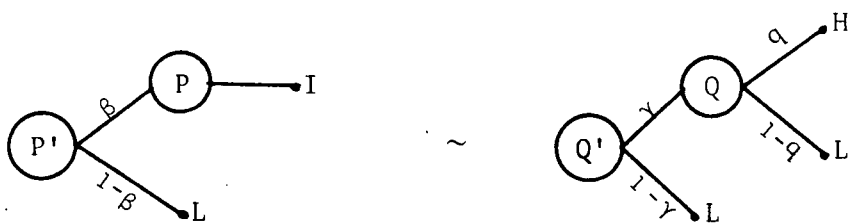


$$v(L) = 0; \quad v(H) = 1;$$

$$v(I) = q \text{ such that } P \sim Q.$$

Having constructed the v -function, we form in Fig. 4.4 the following lotteries to determine the α function.

Figure 4.4: Test of Substitution Axiom



If the decision maker subscribes to the Substitution principle, P' and Q' would be indifferent whenever β equals γ ; since Q is constructed to be indifferent to P . If, however, P' and Q' are indifferent with β and γ

unequal, then Ratio Consistency tells us that

$$\frac{\gamma/1-\gamma}{\beta/1-\beta} = \tau, \text{ a constant.} \quad (4.13)$$

The decision maker's α -function is then given by $\alpha(I)=\tau$ and $\alpha(L)=\alpha(H)=1$, i.e., assign 1 to the α -values of the best (H) and the worst (L) outcomes of the consequence set X. (See the sufficiency proof of Theorem 2.1 in Chapter 2 for more details).

For consequence sets with more than three outcomes, the above procedure can be repeated for the other intermediate outcomes. In the case of an interval of a real line, the v and α functions can be obtained by interpolating among a finite number of measurement points. Just as there are many different ways to measure von Neumann-Morgenstern utility functions, this would also be the case for alpha utility. Since we only skim the issue of assessment here, more work seems to be needed. In the next sub-sections, we obtain conditions for consistency of alpha utility with stochastic dominance, local and global risk aversion.

4.3.3 Stochastic Dominance

Definition 4.1: A distribution G is said to stochastically dominate another distribution F in the first degree, denoted by $G \succeq^1 F$, if

$$G(x) \leq F(x), \quad \forall x \in \mathbb{R}.$$

The above definition of what is usually called *Stochastic Dominance* has its origin in the works of Hadar and Russell (1969), Lehmann (1955) and Hanoch and Levy (1969). It is well known that,

$$G \succeq^1 F \Rightarrow \int_{\mathbb{R}} u dG \geq \int_{\mathbb{R}} u dF$$

for every increasing function u on \mathbb{R} . Therefore, every expected utility decision maker would prefer G to F when G dominates F in the first degree. This is however not necessarily true of every alpha utility decision maker. To the extent that consistency with first degree stochastic dominance is normatively desirable (and, in the context of monetary lotteries, probably descriptively valid), it can be imposed as an additional requirement. The following corollary taken from Chapter 1 provides conditions under which an alpha utility decision maker would be consistent with first degree stochastic dominance.

Corollary 1.4: Suppose α, v are bounded.

$$\begin{aligned} \text{Then } \forall F, G \in \mathcal{D}, F \succeq^1 G &\Rightarrow \Omega(F) \geq \Omega(G) \\ \text{iff } \alpha(x)(v(x) - v(s)) & \hspace{10em} (4.14) \\ \text{is an increasing function } \forall s \in \mathbb{R}. \end{aligned}$$

The functional $\Omega(F)$ refers to the alpha utility representation $\int_{\mathbb{R}} \alpha v dF / \int_{\mathbb{R}} \alpha dF$. It is easy to see that relation (4.14) can be restated,

$$\alpha(x) \cdot (v(x) - \Omega(F)) \quad (4.15)$$

is an increasing function $\forall F \in D$.

We obtain a "familiar" interpretation of (4.15) through the following expository discussion of the necessity proof of Corollary 1.5. Observe that $G \succeq F \Rightarrow F_{\theta} \succeq F_{\theta'}$ if $\theta' > \theta$ where $F_{\theta} \equiv (1-\theta)F + \theta G$. Suppose Ω is consistent with \succeq , then $\Omega(F_{\theta})$ is increasing in θ . Differentiating with respect to θ yields

$$\frac{d}{d\theta} \Omega(F_{\theta}) = \int_{\mathbb{R}} u(x; F_{\theta}) d[G(x) - F(x)] \geq 0 \quad \forall \theta \in (0,1] \quad (4.16)$$

$$\text{where } u(x; F_{\theta}) = \frac{\alpha(x)}{E(\alpha, F_{\theta})} \{v(x) - \Omega(F_{\theta})\}. \quad (4.17)$$

By continuity of $u(x; F_{\theta})$ in θ , it follows that

$$\left. \frac{d}{d\theta} \Omega(F_{\theta}) \right|_{\theta=0+} = \int_{\mathbb{R}} u(x; F) d[G(x) - F(x)] \geq 0. \quad (4.18)$$

Since G is any distribution that dominates F , $u(x; F)$ has to be increasing for every $x \in \mathbb{R}$. The function $u(x; F)$ has the functional-analytic interpretation as a Gâteaux derivative (Luenberger, 1969) of the functional $\Omega(F)$ at F and

$$\left. \frac{d}{d\theta} \Omega(F_{\theta}) \right|_{\theta=0+} = \int_{\mathbb{R}} u(x; F) d(G-F)$$

is interpreted as the Gâteaux differential of Ω at F in the direction of $G-F$. For expected utility, the corresponding Gâteaux derivative and Gâteaux differential are $u(x)$ and $\int_{\mathbb{R}} u(x) d[G(x) - F(x)]$ respectively.

Note that $u(x)$ does not depend on F since the expected utility repre-

sensation $\int_{\mathbb{R}} u dF$ is linear in F . This suggests the term *Lottery Specific Utility Function* (abbrev. LOSUF) for $u(x;F)$.⁵ Corollary 1.5 has the familiar interpretation: Alpha utility is consistent with 1st degree stochastic dominance if its LOSUF is increasing for each F . This condition assumes the following more manageable form if we impose 1st order differentiability on α and v .

$$\alpha v' \geq \max[\alpha' [\bar{v} - v(x)], -\alpha' [v(x) - \underline{v}]] \quad \forall x \in \mathbb{R}, \quad (4.19)$$

$$\text{where } \bar{v} = \lim_{x \rightarrow \infty} v(x) \text{ and } \underline{v} = \lim_{x \rightarrow -\infty} v(x).$$

Alternatively, the above can be stated as:

$$\begin{aligned} \forall x \in \mathbb{R}, \quad (\log \alpha(x))' &\leq \frac{v'(x)}{\bar{v} - v(x)} && \text{if } \alpha'(x) \geq 0 \\ &> -\frac{v'(x)}{v(x) - \underline{v}} && \text{if } \alpha'(x) < 0 \end{aligned} \quad (4.20)$$

The rate of change of $\log \alpha$ is bounded from above and below by $v'(x)/\bar{v} - v(x)$ and $-v'(x)/v(x) - \underline{v}$ respectively. Expression (4.20) can be used directly to find all those α given a certain v that will be consistent with 1st degree stochastic dominance.

⁵ Machina (1980) investigated the properties of a twice Fréchet-differentiable functional $V(F)$ on $D[0,M]$ with respect to the stochastic dominance and global risk aversion partial orders. Fréchet differentiability is a stronger notion than Gâteaux differentiability and does not admit a natural extension of the analysis to the real line. Machina called the first Fréchet derivatives of $V(F)$ local utility functions. The alpha utility functional $\Omega(F)$ when restricted to a compact interval such as $[0,M]$ is Fréchet differentiable so that we can identify a lottery specific utility function with a local utility function.

4.3.4 Global Risk Aversion

Definition 4.2: A distribution G is said to stochastically dominate another distribution F in the second degree,⁶ denoted by $G \stackrel{2}{\succeq} F$,

$$\text{if} \quad \int_{-\infty}^y (G(x) - F(x)) dx \leq 0 \quad \forall y \in \mathbb{R} \quad (4.21)$$

$$\text{and} \quad \int_{-\infty}^{\infty} (G(x) - F(x)) dx = 0. \quad (4.22)$$

When the means associated with the distributions F and G exist, condition (4.22) implies that they are equal. Condition (4.21) requires that for each x , the mean of G truncated at $(-\infty, x]$ is not less than that of F truncated at $(-\infty, x]$. Note also that for all concave u in \mathbb{R} , $F \stackrel{2}{\succeq} G$ if and only if $\int_{\mathbb{R}} u dF \geq \int_{\mathbb{R}} u dG$.

Consequently, an expected utility decision maker with a concave utility function always prefers a distribution F that dominates another distribution G in the second degree. The normative content of 2nd-degree Stochastic Dominance (or Global Risk Aversion) is derived from the idea that a prudent person should be risk averse.

Definition 4.2 extends the definition of increasing risk by Rothschild and Stiglitz (1970) for a compact interval. Similar or related results have been obtained in Blackwell (1951), Hanoch and Levy (1969) and

⁶ Whitmore (1970) defined third-degree stochastic dominance and related it to the Arrow-Pratt index. The general k th-degree stochastic dominance was defined in Chapter 1. However, both definitions will not be treated here.

Strassen (1965).

The following corollary adapted from Chapter 1, which is a restriction of Corollary 1.6 to the case of second-degree stochastic dominance, provides conditions under which an alpha utility decision maker would be consistent with second-degree stochastic dominance.

Corollary 1.6*: Suppose α , v , α' , v' are bounded and continuous,

$$\text{then } \forall F, G \in D, \quad F \overset{2}{\succeq} G \Rightarrow \Omega(F) \succeq \Omega(G) \quad (4.23)$$

$$\text{iff } \{\alpha(x)[v(x)-v(s)]\}' \text{ is a decreasing function} \\ \forall s \in R.$$

Note that condition (4.23) requires the LOSUF $u(x;F)$ to be a concave function for each F . This further confirms our intuition that the LOSUF plays a von Neumann-Morgenstern utility-like role.

Requiring α , v to be second differentiable, the condition (4.23) becomes

$$\forall x \in R, \quad \alpha v'' + 2\alpha' v' \leq \min[\alpha''(\bar{v}-v(x)), -\alpha''(v(x)-\underline{v})]. \quad (4.24)$$

This condition is related to the Arrow-Pratt measure for alpha utility developed in the next subsection.

4.3.5 Local Risk Aversion: The Arrow-Pratt Index

Consider an alpha utility decision maker with assets x . The *Cash Equivalent* $C(x;Z)$ corresponding to a risk Z is given by $M(F_{x+Z}) - x$ which is the difference between the certain asset position $M(F_{x+Z})$ such that the decision maker is indifferent to taking the risk Z and his current asset position x . The risk premium $\pi(x;Z)$ is then defined by

$$\pi(x;Z) = E(Z) - C(x;Z) \quad (4.25)$$

which is the difference between the actuarial value $E(Z)$ of the risk Z and its cash equivalent $C(x;Z)$. Since $x+Z$ and $(x+\mu)+(Z-\mu)$ have the same distribution on final assets, π and C from expression (4.25) have the properties:

$$C(x+\mu;Z-\mu) = C(x;Z) - \mu, \quad \text{and} \quad (4.26)$$

$$\pi(x+\mu;Z-\mu) = \pi(x;Z). \quad (4.27)$$

Following Pratt (1964), we consider the behaviour of $\pi(x;Z)$ for an actuarially fair risk Z as $\sigma_Z \rightarrow 0$, assuming the third absolute central moment of Z is of order $o(\sigma_Z^2)$. Thus,

$$v(x - \pi(x;Z)) = \Omega(F_{x+Z}) = \frac{\int_R \alpha v dF_{x+Z}}{\int_R \alpha dF_{x+Z}}. \quad (4.28)$$

Expanding both sides after cross multiplication, we get

$$\begin{aligned}
& [v(x) - \pi v'(x) + o(\pi^2)] [\alpha(x) + \frac{\sigma_Z^2}{2} \alpha''(x) + o(\sigma_Z^2)] \\
& = \alpha(x) v(x) + \frac{\sigma_Z^2}{2} (\alpha v)''(x) + o(\sigma_Z^2).
\end{aligned} \tag{4.29}$$

This reduces to

$$\pi(x; Z) = \frac{1}{2} \sigma_Z^2 r(x) + o(\sigma_Z^2), \tag{4.30}$$

$$\text{where } r(x) = -\frac{d}{dx} [\log \alpha^2(x) v'(x)]' = -\left(\frac{v''(x)}{v'(x)} + \frac{2\alpha'(x)}{\alpha(x)} \right). \tag{4.31}$$

As in expected utility, the decision maker's risk premium for a small, actuarially neutral risk Z is approximately half the variance times $r(x)$, which, in keeping with precedent, we call the Arrow-Pratt index. When Z is not actuarially fair, we obtain from (4.27)

$$\pi(x; Z) = \frac{1}{2} \sigma_Z^2 r(x + E(Z)) + o(\sigma_Z^2). \tag{4.32}$$

It is straightforward to check that $r(x)$ has the following alternative interpretation in terms of the probability premium (Pratt, 1964; Arrow, 1971),

$$p(x; h) = \frac{1}{2} h r(x) + O(h^2), \tag{4.33}$$

where $\frac{1}{2}(1+p(x; h))$ and $\frac{1}{2}(1-p(x; h))$ are the probabilities of obtaining $x+h$ and $x-h$ respectively, such that the decision maker is indifferent between the status quo x and taking the risk. It is also straightforward to check that $r(x)$ is invariant under the uniqueness transformation (expression (1.12) and (1.13)) for the functions α and v . In particular, it is invariant under an affine transformation for v and a scalar

multiple for α .

It is comforting to note that $r(x)$ has similar local properties as expected utility. Unlike expected utility however, we cannot recover the functions α and v only from the knowledge of $r(x)$ pointwise. We would not therefore expect to be able to characterize in general, global risk propensities of an alpha utility decision maker in terms of his local risk aversion function $r(x)$. In particular, we note from subsection 4.3.4 that global risk aversion in the sense of consistency with second degree stochastic dominance implies that $r(x) \geq 0$ pointwise, but the converse does not in general hold. This result seems intuitively appealing. Consistency with second degree stochastic dominance implies positive risk premium $\pi(x;Z)$ at any asset position x , for any actuarially fair risk Z , which in turn implies that $r(x)$ is positive for actuarially fair infinitesimal risks about x . Even if $r(x)$ is positive pointwise, indicating aversion to infinitesimal actuarially fair risks, it is still possible for an alpha utility decision maker to be risk seeking over some interval. This corresponds to the observation that people purchase insurance and gamble at the same time (Friedman & Savage, 1948; Markowitz, 1952). The implications of this and other empirically observed choice behavior for alpha utility will be considered in the next section; where we use the consistency conditions developed in the last subsections to identify regions of local and global risk aversion.

4.4 DESCRIPTIVE IMPLICATIONS

We showed in the preceding section that alpha utility, like expected utility, is compatible with such normative notions as stochastic dominance, local and global risk aversion. We also illustrated how the constructive proof of our representation theorem furnishes a procedure for the assessment of the alpha utility functions. In this section, we will show that alpha utility is not so general that it has no testable implications, nor is it such a minute departure from expected utility that it is susceptible to the same set of violations.

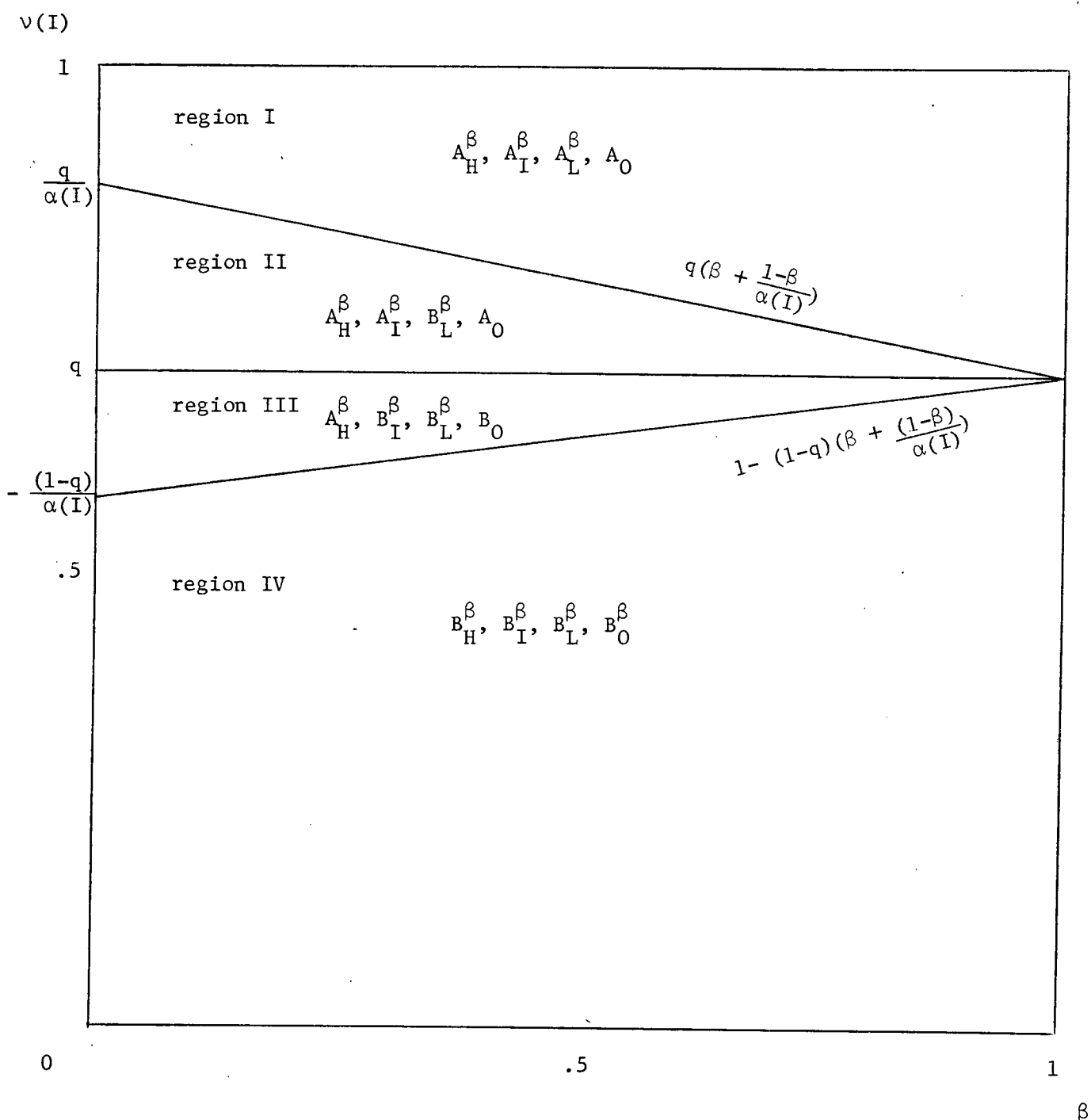
4.4.1 Systematic Violations of the Strong Independence Principle⁷

The empirical findings violating the implications of the Strong Independence principle that stem from the Allais paradox were summarized in table 3.1 in terms of the HILO structure of lotteries developed in section 3.1. We consider here the implications of alpha utility theory for choice patterns for the HILO structure, and derive some testable predictions.

Applying alpha utility theory to the decision choices in Figure 3.4 we obtain,

$$\begin{array}{ll}
 \Omega(A_0) = v(I) & \Omega(B_0) = q \\
 \Omega(A_L^\beta) = \frac{\beta\alpha(I)v(I)}{\beta\alpha(I) + (1-\beta)} & \Omega(B_L^\beta) = \beta q \\
 \Omega(A_I^\beta) = v(I) & \Omega(B_I^\beta) = \frac{\beta q + (1-\beta)\alpha(I)v(I)}{\beta + (1-\beta)\alpha(I)} \\
 \Omega(A_H^\beta) = \frac{\beta\alpha(I)v(I) + (1-\beta)}{\beta\alpha(I) + (1-\beta)} & \Omega(B_H^\beta) = \beta q + (1-\beta)
 \end{array}$$

⁷ The material in this sub-section appeared in Chew and MacCrimmon (1979b).

Figure 4.5: Preference pattern for $\alpha(I) < 1$

Assuming without loss of generality that $\alpha(L)=\alpha(H)=1$, (cf. Theorem 2.1), we obtain the following inequalities corresponding to preferences in each decision box.

$$A_O > B_O \Leftrightarrow v(I) > q \Leftrightarrow A_I^\beta > B_I^\beta \quad (4.34)$$

$$A_L^\beta > B_L^\beta \Leftrightarrow v(I) > q(\beta + \frac{(1-\beta)}{\alpha(I)}) \quad (4.35)$$

$$A_H^\beta > B_H^\beta \Leftrightarrow v(I) > 1-(1-q)(\beta + \frac{(1-\beta)}{\alpha(I)}) \quad (4.36)$$

The inequalities (4.34), (4.35) and (4.36) are plotted in Figure 4.5 for the case in which $\alpha(I) < 1$. Note that the four regions, denoted as regions I-IV, correspond to four distinct choice patterns. These, along with the choice patterns for the case of $\alpha(I) > 1$, are summarized in Table 4.1.

Table 4.1

Allowable Choice Patterns Under Alpha Utility Theory

| $\alpha(I)$ Region | < 1 | Expected Utility = 1 | > 1 |
|-----------------------|----------------------|----------------------------|----------------------|
| I | A_H, A_I, A_L, A_O | | |
| II | A_H, A_I, B_L, A_O | Regions do not exist | B_H, A_I, A_L, A_O |
| III | A_H, B_I, B_L, B_O | | B_H, B_I, A_L, B_O |
| IV | B_H, B_I, B_L, B_O | | |

Apart from the patterns corresponding to regions I and IV (the only ones consistent with expected utility) alpha utility theory allows for 4 out of 14 additional patterns. These are given by the entries

under the case $\alpha(I) < 1$ and $\alpha(I) > 1$ for regions II and III. The region II and region III patterns under $\alpha(I) > 1$ have not been reported in the literature (cf. Table 3.1). On the other hand, all the empirical findings to date of violations of expected utility correspond to either of the region II and III patterns with $\alpha(I) < 1$. In particular, both the standard Allais paradox (i.e. A_I^β, B_L^β) and the Allais ratio paradox (i.e. B_L^β, A_O) occur in region II. The existence of region III also has some empirical support; note (Table 3.1) that both the A_H^β, B_O and $A_H^{\beta 1}, B_H^{\beta 2}$ violations have been reported.

Before we continue with further implications based on the allowable patterns of choice, we can gain some intuition about alpha utility theory in relation to the HILO structure by examining, in Figure 4.5, the effects of changes in the parameters β and $v(I)$ on the resulting pattern of choices. Looking first at changes in β , we see that if consequence I is sufficiently attractive (i.e., $v(I) > q/\alpha(I)$) then the choice will be A alternative and will be unaffected by changes in β . This indicates a basis for choice that may be called the *security effect*. Correspondingly, if I is unattractive (i.e., $v(I) < 1 - (1-q)/\alpha(I)$), then no change in β will induce a shift away from the B alternative. This may be called a *nothing-to-lose effect*. Note that both above regions are the only regions consistent with the substitution principle.

If I is somewhat attractive (i.e., $v(I)$ between q and $q/\alpha(I)$) then decreasing β from 1 will cause a switch from A_L^β to B_L^β . The smaller value of β acts as a dilution probability to narrow the perceived gap between A_L^β and B_L^β , due to the attractiveness of I, until finally the attractiveness is sufficiently diluted to cause a switch to B_L^β . This

is the *dilution effect*. The special case for the dilution from $A_L^{\beta_1}$ to $B_L^{\beta_2}$, (where $\beta_1 = 1.0$ and $\beta_2 \ll 1.0$) has been called a certainty effect (Kahneman and Tversky, 1979). When I is somewhat unattractive (i.e., $v(I)$ is between $1 - (1-q)/\alpha(I)$ and q) a decrease in β will cause a switch from B_H^β to A_H^β . This is a reverse dilution effect.

In addition to studying the effect of changes in β for given values of $v(I)$, fresh insight can be gained by examining the effect of changes in $v(I)$ for given levels of β . For $\beta = 1$, we no longer have a compound structure and so revert to the simple A_0, B_0 choice which depends on whether $v(I) > q$. At low and intermediate levels of β , the regions II and III, giving rise to the paradoxical choices can be quite large. Note that as $v(I)$ decreases from being very attractive, at some point the A_L^β choice will switch to B_L^β . Then as $v(I)$ drops below q , the choices in the I and 0 cases change to B_I^β and B_0 respectively. Finally, as $v(I)$ drops lower (i.e., below $q(\beta + (1-\beta)/\alpha(I))$), A_H^β changes to B_H^β .

Two main implications should be noted from these observations. First, alpha theory can describe a richer set of preferences than can expected utility theory. Specifically, it covers the Allais-type preferences (I-L, L-0, and L-L) and other observed violations (H-0 and H-H). It allows for a dependence on the values of the parameter and hence captures the observed *dilution effect*.

On the other hand, it is not so general that it can describe any preferences. Only 6 of the possible 16 preference patterns of the HILO structure are consistent with alpha utility theory (and only 4 patterns for the empirically supported $\alpha(I) < 1$ case). As Figure 4.5

suggests, our theory makes very specific predictions about the preference patterns and the way they change as the parameters β and $v(I)$ changes. In particular, monotonicity requires that preferences for the O and I cases agree completely, i.e., $A_O \succ B_O \iff A_I \succ B_I$. More interestingly, note that the standard Allais paradox (A_I, B_L) occurs if, and only if, the Allais ratio paradox (B_L, A_O) occurs. Both new regions of permissible choice (i.e., regions II and III) provide for A_H^β and B_L^β ; hence this previously unreported case would seem to be a prime candidate for a new "paradox". Further, note that as $v(I)$ decreases, the switch from A^β to B^β occurs first for the L case, then for the I and O cases, and finally for the H case. All these implications from alpha utility theory are empirically testable.

4.4.2 Stochastic Dominance

Consistency with stochastic dominance is an intrinsic property of expected utility. This is not the case with alpha utility. To the extent that it is prescriptively desirable and descriptively valid (see Chapter 5 for an example of a potential violation of stochastic dominance in the context of income distributions), we can restrict the alpha utility functions considered to those that do not violate stochastic dominance via the following consistency conditions taken from Corollary 1.4.

$$\begin{aligned} \forall x \in \mathbb{R}, [\log \alpha(x)]' &\leq - [\log (\bar{v} - v(x))]' \text{ if } \alpha'(x) \geq 0 \\ &\geq - [\log (v(x) - \underline{v})]' \text{ if } \alpha'(x) < 0. \end{aligned}$$

$$\text{where } \bar{v} = \lim_{x \rightarrow \infty} v(x) \text{ and } \underline{v} = \lim_{x \rightarrow -\infty} v(x) \quad (4.37)$$

These conditions are depicted graphically in Figure 4.6 for a bounded v normalized to $\underline{v}=0$ and $\bar{v}=1$. The slope of $\log \alpha$ is bounded by the slope of $-\log (1-v)$ from above and $-\log v$ from below. The α function that is compatible with (4.37) is then recovered from the graph of $\log \alpha$. The α function considered has a dent in the middle and increases in both the positive and the negative directions in order not to exclude the possibility of region II and region III preferences in Table 4.1. (Recall that the condition for the existence of region II and III preference is that $\alpha(I) < q\alpha(H) + (1-q)\alpha(L)$). The limiting behavior of α is bounded above by $\frac{1}{1-\underline{v}}$ and below by $\frac{1}{\underline{v}}$.

It is interesting to note that after satisfying stochastic dominance and also the "Allais" type preference for the HILO structure, we are still left with a fairly large class of α functions. We investigate in the next sub-section, the risk propensities of the α

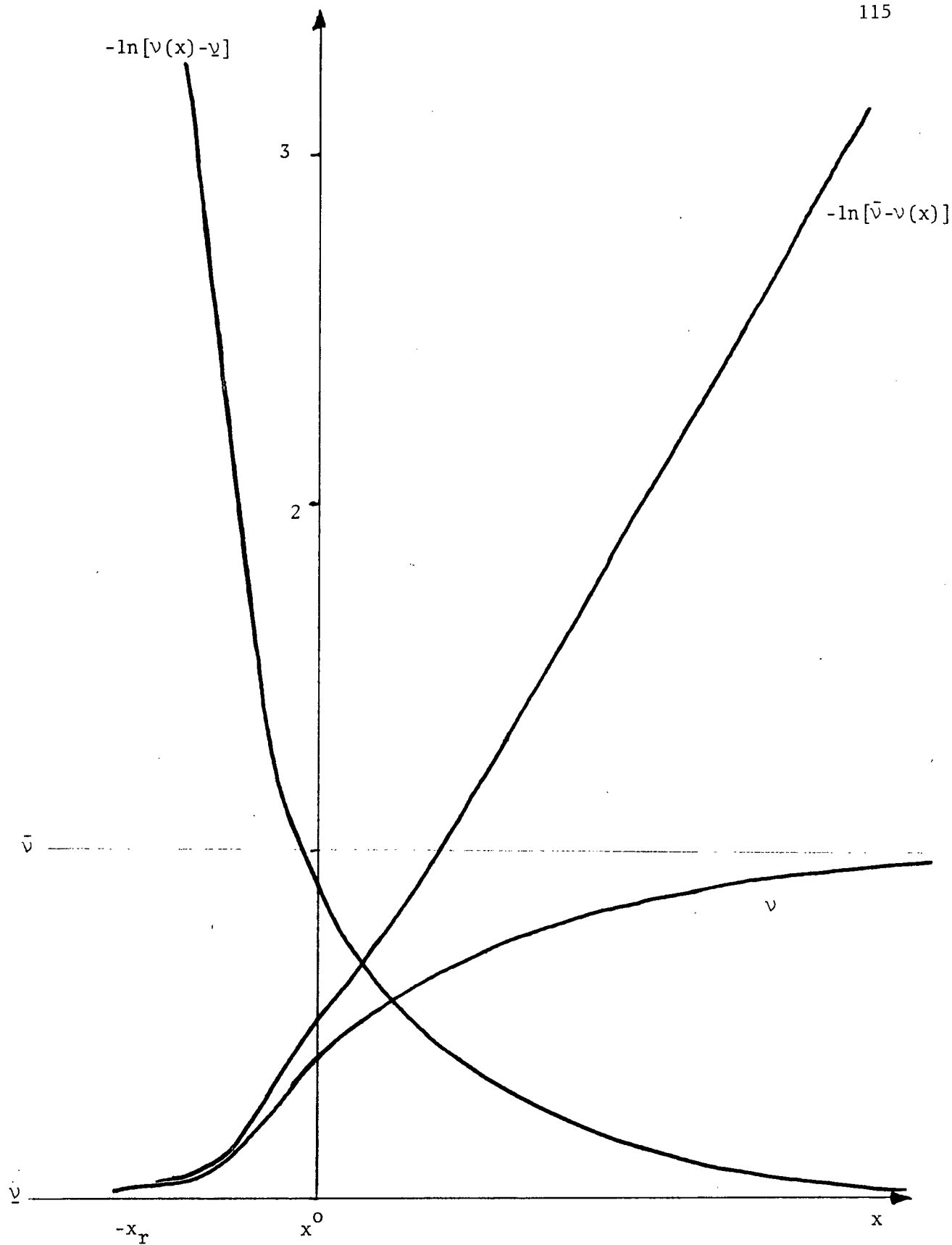


Fig. 4.6: Consistency Conditions for Stochastic Dominance

and v functions of Figure 4.6.

4.4.2 Local and Global Risk Properties: Concurrence of Risk Averting and Risk Seeking Behavior

Local risk aversion in the sense that the Arrow-Pratt index (see sub-section 4.3.5) is nonpositive corresponds to the observation that people tend to avoid taking a small gamble in favor of its expected return. For expected utility, local risk aversion pointwise is equivalent to having a concave von Neumann-Morgenstern utility function which is in turn equivalent to global risk aversion. Thus, the behaviorally plausible local risk aversion hypothesis is not compatible with any concurrent risk seeking behavior, e.g., the purchase of a lottery ticket. Alpha utility does not share the above difficulty since local risk aversion is necessary but not sufficient for global risk aversion. The condition for local risk aversion, namely that the Arrow-Pratt index $(= -\frac{v''(x)}{v'(x)} - \frac{2\alpha'(x)}{\alpha(x)})$ is nonnegative, is:

$$(\log \alpha(x))' \leq -\frac{1}{2}(\log v'(x))'. \quad (4.38)$$

This is depicted graphically in Figure 4.7 for the v function of Figure 4.6. Note that a constant α (which corresponds to expected utility) is not admissible. The function $\log \alpha$ has to decrease sufficiently rapidly near the ruin point $-x_r$ in order to correct for the convexity of v at the same region.

At this point, we have a pair of functions (α, v) that satisfy stochastic dominance, exhibit local risk aversion pointwise, and are compatible with the Allais type preference for the HILO structure. It remains to check whether the functions have the correct global risk properties that correspond to actual choice behavior. First, we establish

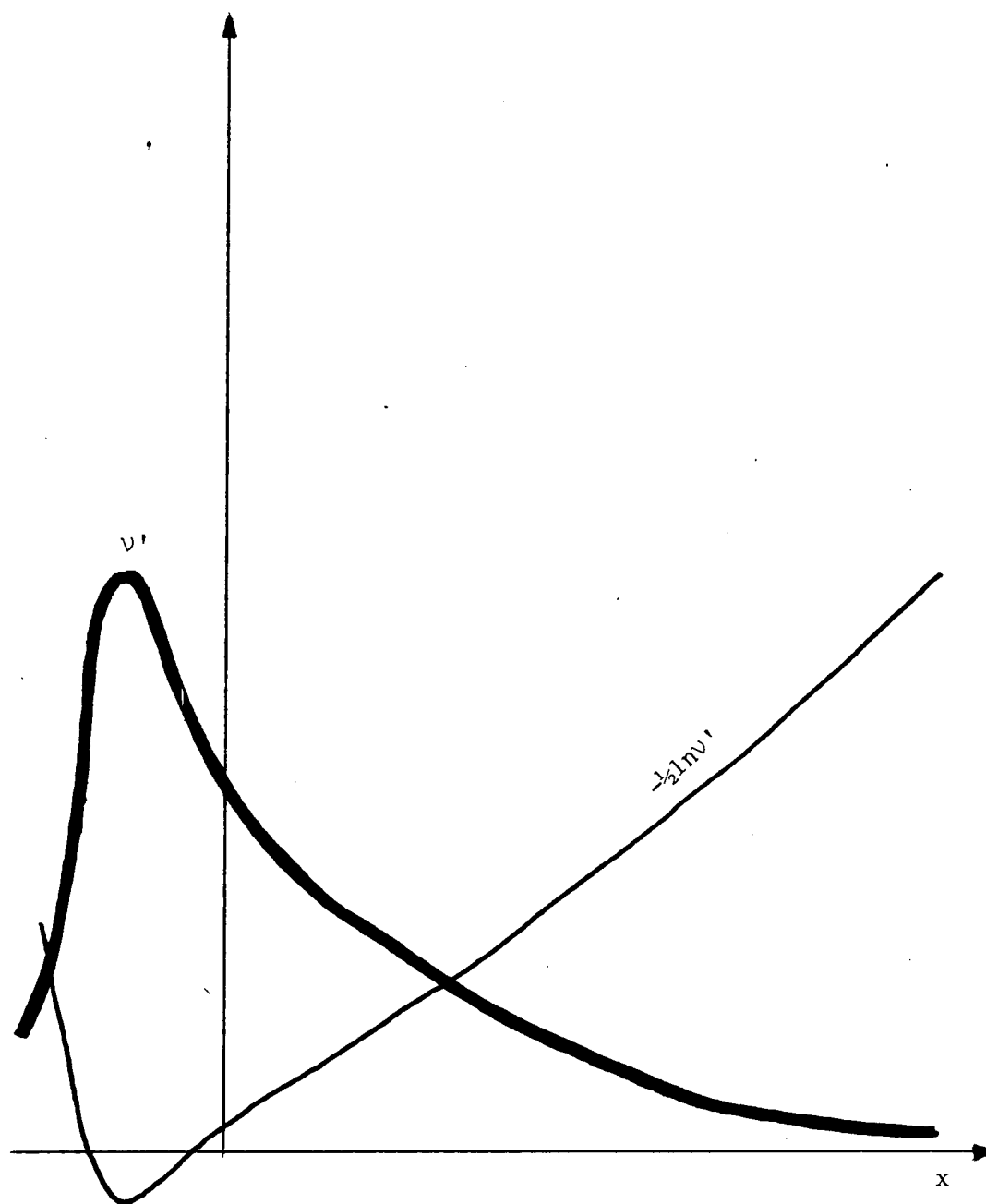


Fig. 4.7: Conditions for Local Risk Aversion

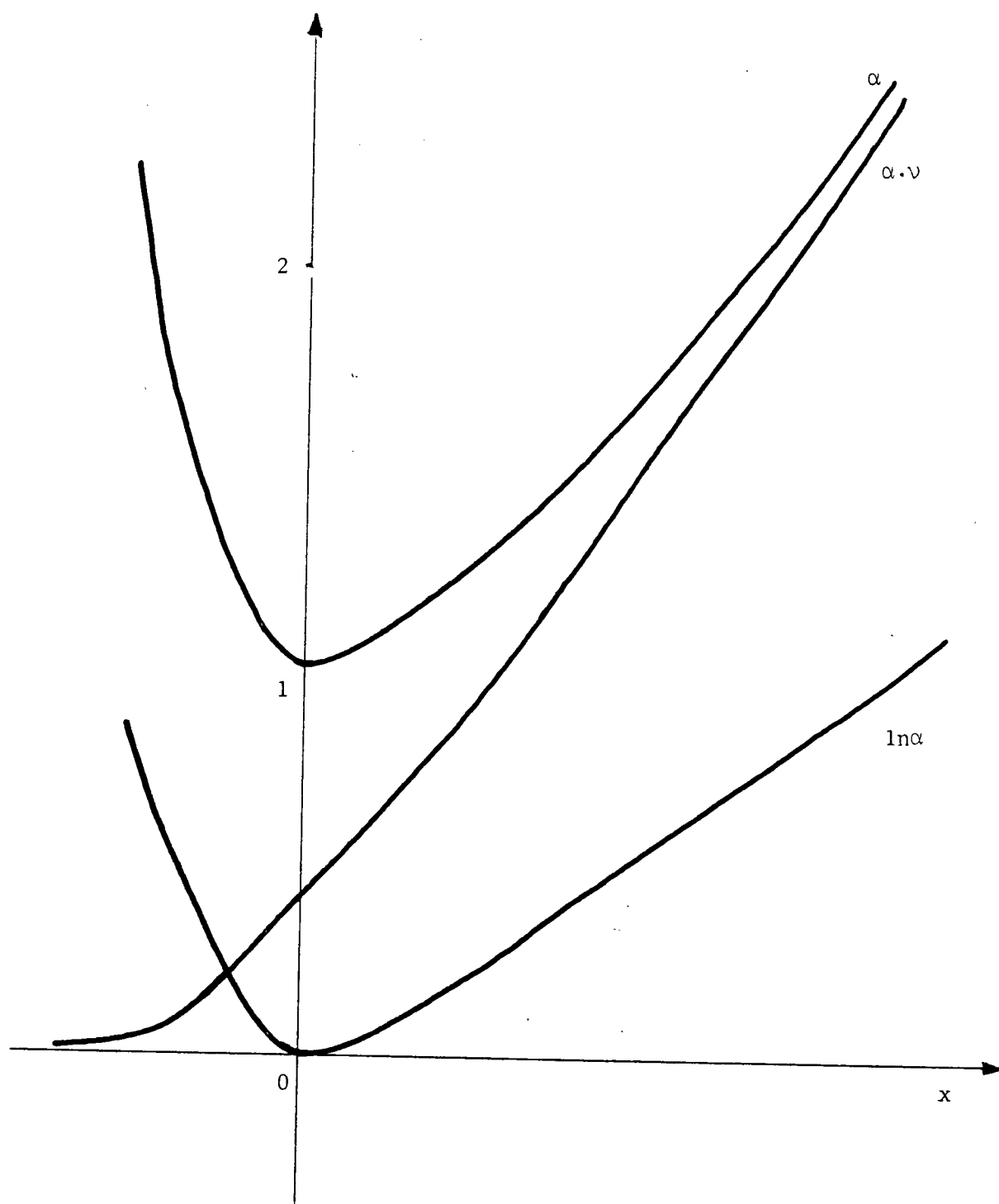


Fig. 4.8: An admissible alpha function

that our α and v functions describe risk seeking behavior by showing that they do not satisfy the conditions for global risk aversion:

$$\begin{aligned} \forall x \in \mathbb{R}, (\alpha(x) \cdot v(x))'' &\leq 0 && \text{if } \alpha''(x) \geq 0 \\ &\leq \alpha''(x) && \text{if } \alpha''(x) < 0 \end{aligned} \quad (4.39)$$

Since the α function considered is convex, it suffices to observe in Figure 4.8 that the product $\alpha \cdot v$ admits a convex region near the ruin point. This means that there is a concurrence of risk seeking and risk averting behavior for that region. This corresponds to the prevalence of risk seeking behavior for losses observed by Markowitz (1952) and recently by Kahneman and Tversky (1979). However, the local risk aversion hypothesis (see Figure 4.6) rules out the possibility of global risk proneness so that the alpha utility decision maker will at the same time have the opposite risk averse propensity for some other gambles.

Finally, we consider the mutual incompatibility of different procedures for the assessment of a von Neumann-Morgenstern utility function. Suppose we apply the certainty equivalent method to the alpha utility decision maker considered, with status quo $x=0$ and some substantial gain amount x_g as the endpoints of lottery B in Figure 3.5. The pairs of measurements $(x_c, u_c(x_c))$ that corresponds to indifference between getting x_c for sure in lottery A, and getting x_g with $u_c(x_c)$ chance and getting 0 with $1-u_c(x_c)$ chance are related by the following expression:

$$v(x_c) = \frac{\alpha(x_g)u_c(x_c)v(x_g) + \alpha(0)[1-u_c(x_c)]v(0)}{\alpha(x_g)u_c(x_c) + \alpha(0)[1-u_c(x_c)]} \quad (4.40)$$

After rearranging, we get:

$$u_c(x) = \frac{\alpha(0)[v(x)-v(0)]}{\alpha(0)[v(x)-v(0)] + \alpha(x_g)[v(x_g)-v(x)]} \quad (4.41)$$

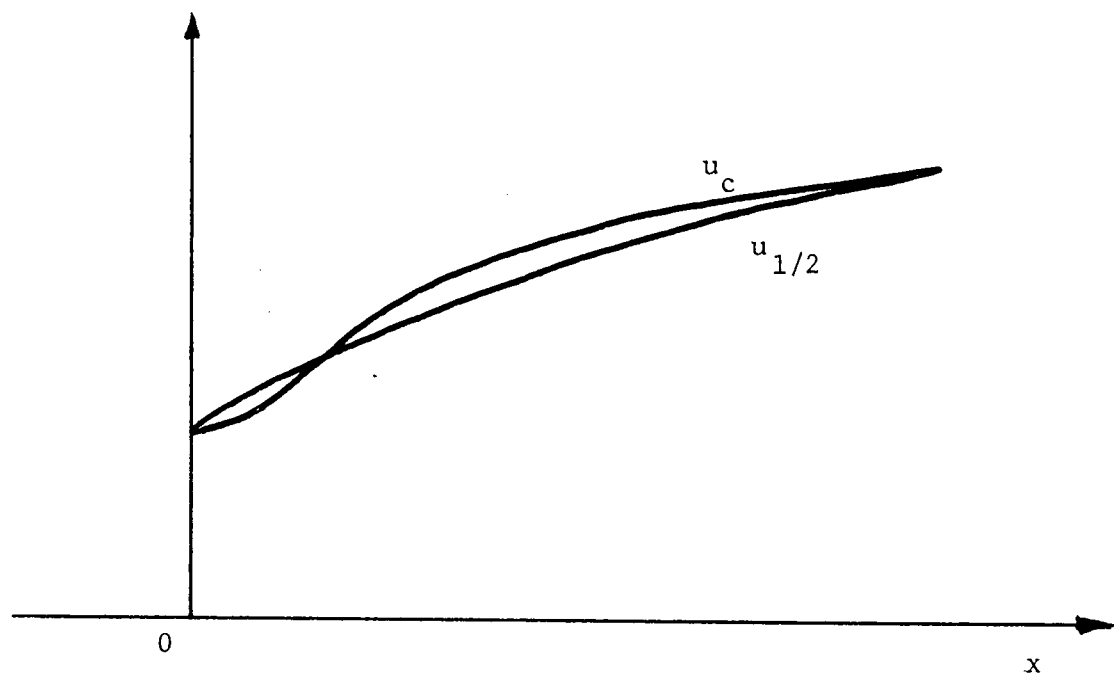


Fig. 4.9: A pair of u_c and $u_{1/2}$ derived from an alpha utility decision maker

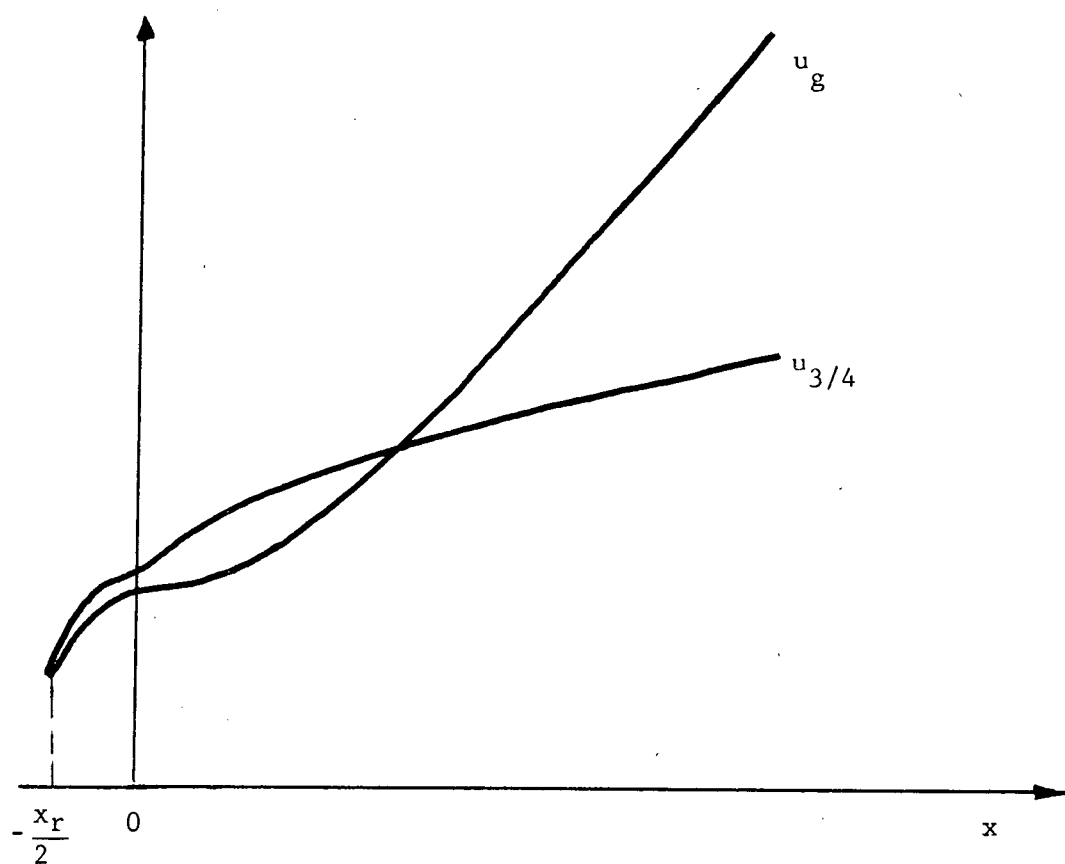


Fig. 4.10: A pair of u_g and $u_{3/4}$ derived from an alpha utility decision maker

Alternatively, if we fix a loss amount at half the ruin position $-x_r$ and use $x=0$ as the intermediate amount, we can apply the gain equivalent method to obtain u_g below:

$$v(0) = \frac{\alpha(x_g)p_g(x_g)v(x_g) + \alpha(-\frac{x_r}{2})[1-p_g(x_g)]v(-\frac{x_r}{2})}{\alpha(x_g)p_g(x_g) + \alpha(-\frac{x_r}{2})[1-p_g(x_g)]}.$$

Hence,

$$u_g(x) = \frac{1}{p_g(x)} = \frac{\alpha(x)}{\alpha(-\frac{x_r}{2})} \left[\frac{v(x)-v(0)}{v(0)-v(-\frac{x_r}{2})} \right] + 1. \quad (4.43)$$

Finally, we consider the chaining method. With $x^0=0$ and $x_g=x_i$ as the endpoint, and $p^0=\frac{1}{2}$ as the probability parameter, we obtain the sequence (x_1, x_2, \dots) based on the following relation.

$$v(x_{i+1}) = \frac{\alpha(x_i)v(x_i) + \alpha(0)v(0)}{\alpha(x_i) + \alpha(0)} \quad (4.44)$$

We have thus determined $u_{1/2}$ on the sequence of points (x_1, x_2, \dots) given by expression (4.44). Changing the probability parameter to $p^0=3/4$ and the lower endpoint to $x_1^0=-\frac{x_r}{2}$, we obtain:

$$v(x_{i+1}) = \frac{3\alpha(x_i)v(x_i) + \alpha(-\frac{x_r}{2})v(-\frac{x_r}{2})}{3\alpha(x_i) + \alpha(0)} \quad (4.45)$$

As before, we have determined $u_{3/4}$ on the sequence of points (x_1, x_2, \dots) determined by expression (4.45).

Figure 4.9 shows the graphs of u_c and $u_{1/2}$ with the same axes for ease of comparison with Allais' results displayed in Figure 3.6.

Similarly, the curves of $u_{3/4}$ and u_g are plotted in Figure 4.10 for comparison with the results of MacCrimmon et. al. (1972) in Figure 3.7.

The rather striking fit between the theoretical prediction and the empirical measurement supports the validity of alpha utility theory;

thus providing a "rational" explanation for an otherwise puzzling phenomena.

4.4.4 Some Problems with Problem Representation

In this final subsection, we comment briefly on how alpha utility handles the two difficulties with problem representation discussed in section 3 of Chapter 3. Our position on the question of whether alpha utility functions should be defined on gains and losses relative to a customary wealth level or in terms of final asset positions is the same as that of expected utility: it depends on the particular application. The latter should be adopted if we are interested in modeling the choice behavior of the "economic" man or if a decision analyst is helping someone who professes his belief in integrating possible outcomes into his wealth position prior to evaluation (Kahneman and Tversky called this the asset integration position). On the other hand, the former position comes in handy if we want to describe the actual choice behavior of people who do not conform to the asset integration position.

On the other question of whether a two-stage lottery is equivalent to its single-stage decomposition. Kahneman and Tversky showed that for a particular two-stage lottery which is closely related to their version of the Allais paradox (cf. the 0-L case in Table 3.1), a problem description that focuses the subject's attention on a conditional comparison between one of the branches can elicit a majority preference that is the reverse of what would be the case if the two-stage lottery is stated in terms of its single-stage equivalent.

Expected utility cannot be consistent with this phenomena because the strong independence principle or the sure-thing principle dictates that the direction of the conditional preference must be the same as that between the overall preference. Alpha utility does not share this difficulty, since the conditional preference for A_0 against B_0 and the

preference for B_L against A_L is a special case of the HILO choice pattern (Table 4.1) considered in sub-section 4.4.1.

The phenomena considered have implications for studies in decision making in general. Apart from situations where a clear normative position dictates which is the 'correct' problem representation, the lesson seems to be that one should be sensitive to the context associated with a particular choice situation. People appear to use different schemes to represent the same choices, resulting in apparent inconsistency if the rankings among alternatives are different for different problem representations.

4.5 CRITIQUE OF ALLAIS' THEORY AND PROSPECT THEORY

Having developed alpha utility theory in the preceding sections, we consider in this section two other alternatives to expected utility that have attracted significant interest.

4.5.1 Allais' Theory

Allais (1953) assumed the existence of a functional $V(F)$ that represents preference among probability distributions. Such an approach has the immediate implications that preference thus represented satisfies several standard properties of expected utility: completeness, transitivity, combination and composition.

Allais further assumed that preference shares another property of expected utility, called consistency with stochastic dominance. This assumption restricts the preference functional $V(F)$ to those that increase when the underlying distribution increases in the stochastic dominance sense. For finite lotteries of the form $F \equiv \sum_{i=1}^n p_i \delta_{x_i}$, this condition has the following simple form.

$$V\left(\sum_{i=1}^n p_i \delta_{x_i}\right) \text{ increases in each } x_i. \quad (4.46)$$

Unlike expected utility, the theory outlined above is not an axiomatic theory in the sense that the existence of the representation $V(F)$ is asserted rather than being a consequence of the assumed properties of the underlying preference. Note also that the adoption of a more general representation is traded against the convenience of having a simple von Neumann-Morgenstern utility function which captures our intuition about diminishing marginal utility and offers a simple characterization of risk proneness (aversion) via the convexity (concavity)

of the utility function.

In order to obtain a utility-like function without resorting to expected utility, Allais revived the Frisch (1926) notion of quaternary preference among intervals of wealth and asserted the existence of a cardinal utility of wealth he termed psychological value (denoted by \bar{s}). In other words, getting \$100 at status quo is "better" than getting \$100 after you have just received \$1000 if the difference in psychological value from status quo to getting \$100, $\bar{s}(\$100) - \bar{s}(\$0)$, is greater than the corresponding difference going from \$1000 to \$1100, $\bar{s}(\$1100) - \bar{s}(\$1000)$. According to Allais, a choice agent's preference depends on the mathematical expectation (first moment), the dispersion (second moment) and in general the shape of the probability distribution of psychological values. This led him to assert that the preference functional $V(F)$ evaluated at a distribution F can be stated in terms of some functional $h(F_{\bar{s}})$ of the distribution of psychological values, $F_{\bar{s}}$, as follows:

$$V(F) = h(F_{\bar{s}}). \quad (4.47)$$

For a finite lottery, $F (\equiv \sum_{i=1}^n p_i \delta_{x_i})$, the corresponding $F_{\bar{s}}$ is given by:

$$F_{\bar{s}} \equiv \sum_{i=1}^n p_i \delta_{\bar{s}(x_i)}. \quad (4.48)$$

After rescaling, we obtained the functional \bar{h} below.

$$\bar{h}(\delta_{\bar{s}(x)}) = \bar{s}(x). \quad (4.49)$$

An important property of \bar{h} , which we will derive shortly, was however not noted by Allais. Since \bar{s} is an interval scale, the functional \bar{h} must represent the same preference under an affine transformation for \bar{s} .

Consider a lottery F and its certainty equivalent $M(F)$. It follows that

$$\bar{h}(F_{\bar{s}}) = \bar{h}(\delta_{\bar{s}(M(F))}) = \bar{s}(M(F)). \quad (4.50)$$

Under the affine transformation $a\bar{s} + b$, where $a > 0$,

$$\begin{aligned}\bar{h}(F_{a\bar{s}+b}) &= \bar{h}(\delta_{a\bar{s}(M(F))+b}) \\ &= a\bar{s}(M(F))+b \\ &= a\bar{h}(F_{\bar{s}})+b.\end{aligned}\quad (4.51)$$

It turns out that the above property of \bar{h} enables us to draw very specific conclusions about the admissible functional forms of \bar{h} . Consider a uniform lottery $F_{\bar{s}} \equiv \frac{1}{N} \sum_{i=1}^N \delta_{\bar{s}(x_i)}$. Applying Aczel's theorem (Aczel, 1966, p. 236) to \bar{h} evaluated at $F_{\bar{s}}$ yields:

$$\begin{aligned}\bar{h}\left(\frac{1}{N} \sum_{i=1}^N \delta_{\bar{s}(x_i)}\right) &\equiv \mu + \sigma g_N\left(\frac{\bar{s}(x_1)-\mu}{\sigma}, \dots, \frac{\bar{s}(x_N)-\mu}{\sigma}\right) \\ \text{where } \mu &= \frac{1}{N} \sum_{i=1}^N \bar{s}(x_i), \quad \sigma^2 = \frac{1}{N} \sum_{i=1}^N (\bar{s}(x_i)-\mu)^2, \\ \text{and } g_N &\text{ is an arbitrary symmetric function,} \\ \text{provided } \sigma &> 0.\end{aligned}\quad (4.52)$$

Define the functional g on the space of uniform distributions as follows:

$$g\left(\frac{1}{N} \sum_{i=1}^N \delta_{y_i}\right) = g_N(y_1, \dots, y_N). \quad (4.53)$$

It follows that

$$\bar{h}\left(\frac{1}{N} \sum_{i=1}^N \delta_{\bar{s}(x_i)}\right) \equiv \mu + \sigma g\left(\frac{1}{N} \sum_{i=1}^N \delta_{\frac{\bar{s}(x_i)-\mu}{\sigma}}\right). \quad (4.54)$$

Note that the above result can be applied to finite lotteries with rational probabilities by taking N to be the least common denominator.

Hence

$$\text{Lemma 4.1: } \bar{h}(F_{\bar{s}}) \equiv \mu + \sigma g\left(F_{\frac{\bar{s}-\mu}{\sigma}}\right) \quad (4.55)$$

where $F_{\bar{s}} \equiv \sum_{i=1}^n p_i \delta_{\bar{s}(x_i)}$, with p_i rational,

$$\mu = \sum_{i=1}^n p_i \bar{s}(x_i), \quad \sigma^2 = \sum_{i=1}^n p_i [\bar{s}(x_i)-\mu]^2,$$

and g an arbitrary functional on the space of finite distributions with rational probability weights.

Without getting into details, we remark that the extension of Lemma 4.1 to remove the restriction of rational weights is straight forward. We need \bar{h} to be continuous with respect to weak convergence (cf. Chapter 2 Axiom M4), and then note that any finite distribution is the weak limit of a sequence of finite distributions with rational weights. This argument can be extended to include the case of distributions with compact supports since they are weak limits of sequences of finite distributions. Further extension to the case of non-compact support can be obtained through an assumption that is similar to Axiom M5 of Chapter 2.

In a recent paper, Allais (1979) considered the following functional form for \bar{h} .

$$\bar{h}(F_{\bar{s}}) = \mu + w(F_{\bar{s}-\mu}) \quad (4.56)$$

He then expressed w in terms of the normalized central moments of

$F_{\bar{s}-\mu}$ such that

$$w(F_{\bar{s}-\mu}) = f\left(\frac{1}{2} \frac{\sigma^2}{\mu^2}, \frac{1}{3!} \frac{m_3}{\mu^3}, \dots\right) \quad (4.57)$$

$$\text{where } m_n = \int_{\mathbb{R}} (\bar{s}-\mu)^n dF.$$

The resultant expression for \bar{h} ,

$$\bar{h}(F_{\bar{s}}) = \mu + f\left(\frac{1}{2} \frac{\sigma^2}{\mu^2}, \frac{1}{3} \frac{m_3}{\mu^3}, \dots\right),$$

however does not satisfy property (4.51). Therefore it is not admissible.

The idea of representing a distribution via its moments can be applied to the standardized distribution $F_{\frac{\bar{s}-\mu}{\sigma}}$ in expression (4.55). We define

the function f in the following way:

$$f(\bar{m}_3, \bar{m}_4, \dots, \bar{m}_n, \dots) = g\left(F_{\frac{\bar{s}-\mu}{\sigma}}\right) \quad (4.58)$$

where \bar{m}_n is the n th-moment of $F_{\frac{\bar{s}-\mu}{\sigma}}$.

Note that the value of f is not affected by an affine transformation of \bar{s} .

A key idea in Allais' criticism of expected utility is that it neglects the higher moments of the psychological value \bar{s} . Presumably, the more moments we include, the closer does the resulting \bar{h} approximate actual preference. As a first step, one is tempted to consider only the first two moments. The general form of \bar{h} for this case is obtained by setting the function f in (4.58) equal to some constant λ , i.e.,

$$\bar{h}(F_{\bar{s}}) = \mu + \lambda \sigma. \quad (4.59)$$

It is however easy to see that \bar{h} above with a nonzero λ violates the assumption of consistency with stochastic dominance. Consider

$F_{\bar{s}} \equiv p\delta_{\bar{s}(x)} + (1-p)\delta_{\bar{s}(y)}$, with $\bar{s}(x) > \bar{s}(y)$. Differentiating $\bar{h}(p\delta_{\bar{s}(x)} + (1-p)\delta_{\bar{s}(y)})$ with respect to p yields:

$$\begin{aligned} \frac{\partial}{\partial p} \{p\bar{s}(x) + (1-p)\bar{s}(y) + \lambda [p^{\frac{1}{2}}(1-p)^{\frac{1}{2}}(\bar{s}(x) - \bar{s}(y))]\} \\ = (\bar{s}(x) - \bar{s}(y)) \left[1 + \frac{\lambda}{2}(1 - 2p/p^{\frac{1}{2}}(1-p)^{\frac{1}{2}})\right]. \end{aligned} \quad (4.60)$$

This derivative is negative for any negative (positive) value of λ , when p is sufficiently close to zero (one). Hence, \bar{h} restricted to the first two moments of \bar{s} cannot be consistent with stochastic dominance even for lotteries with only two outcomes. We are compelled to include the higher moments if we want to get away from expected utility but stay within Allais' framework. The most general \bar{h} that depends on the first three moments is given by

$$\bar{h}(F_{\bar{s}}) = \mu + \sigma f(m_3/\sigma^3). \quad (4.61)$$

A linear approximation of f by $\lambda + \mu m_3/\sigma^3$ produces:

$$\bar{h}(F_{\bar{s}}) = \mu + \lambda\sigma + \mu m_3/\sigma^2. \quad (4.62)$$

This is essentially the same expression adopted by Hagen (1979). Hagen showed that the Allais type choice behavior is compatible with a positive μ as long as the magnitude of λ is not too large. The positive skewness dependence matches our intuition about people's preference evident in the prevalence of lottery ticket purchase. The question of whether the above form of \bar{h} is consistent with stochastic dominance remains. Consider again the two-outcome lottery $F_{\bar{s}} \equiv p\delta_{\bar{s}(x)} + (1-p)\delta_{\bar{s}(y)}$ with $\bar{s}(x)$ greater than $\bar{s}(y)$. The functional \bar{h} evaluated at $F_{\bar{s}}$ gives:

$$\begin{aligned} & \bar{h}(p\delta_{\bar{s}(x)} + (1-p)\delta_{\bar{s}(y)}) \\ &= \bar{s}(y) + \{p + \lambda[p(1-p)]^{1/2} + \mu[p^2 + (1-p)^2]\} [\bar{s}(x) - \bar{s}(y)]. \end{aligned} \quad (4.63)$$

In order that the derivatives of \bar{h} with respect to p is always positive, we again require λ to be zero. The dependence of \bar{h} on σ is then subsumed in the denominator of the m_3/σ^2 term. The resultant \bar{h} without the $\lambda\sigma$ term is given by:

$$\bar{h}(p\delta_{\bar{s}(x)} + (1-p)\delta_{\bar{s}(y)}) = \bar{s}(y) + \{p + \mu[p^2 + (1-p)^2]\} [\bar{s}(x) - \bar{s}(y)]. \quad (4.64)$$

As p tends to one, \bar{h} converges to $\bar{s}(x) + \mu [\bar{s}(x) - \bar{s}(y)]$. Yet, when p equals one, \bar{h} is equal to $\bar{s}(x)$ which is less than $\bar{s}(x) + \mu [\bar{s}(x) - \bar{s}(y)]$. This implies that the lottery $p\delta_x + (1-p)\delta_y$ is preferred to getting the higher amount x for sure when p is sufficiently close to unity. The conclusion that \bar{h} violates stochastic dominance still holds for negative μ if we consider the behavior of \bar{h} as p tends to zero.

Instead of extending our analysis to include higher moments, we shall pause to take stock of what we learned. In place of the expectation of a utility function, Allais asserted that preference is represented by

a functional that depends on the first moment, the dispersion and in general, the shape of the distribution of a psychological value function, which is obtained from comparison among hypothetical changes in wealth position. Based on the property of the psychological value function as an interval scale, we obtained a restriction on the class of admissible preference functionals. It turns out that the particular form of dependence on the moments of the psychological value considered by Allais is not admissible. Next, we considered those admissible functionals that depend only on the first two moments. The resultant functional form, the sum of the first moment and the standard deviation scaled by some constant, is shown to be inconsistent with stochastic dominance. Extending the analysis to include the third moment, we obtained the functional form that Hagen (1979) considered in a recent paper. This functional form is however again shown to violate stochastic dominance.

It is not known whether the problem with stochastic dominance can be averted by incorporating even higher moments. The difficulty with the first three moments suggests however that the psychological value assumption would not lead to a "clean" way to characterize preferences that expected utility fails to capture. In the next sub-section, we introduce prospect theory, which represents a different approach to the problem of descriptive validity of expected utility first identified by Allais via the famed Allais paradox.

4.5.2 Prospect Theory

Prospect theory, developed by Kahneman and Tversky (1979), distinguishes two phases in the choice process: an editing phase in which problem representation rules called editing operations are applied to the offered prospects followed by an evaluation phase. We first introduce these editing operations before discussing their relation to the form of the evaluation functions.

Coding. The perception of outcomes in terms of gains and losses relative to some reference wealth level. This is usually taken to be the current asset position, in which case gains and losses are the actual amounts to be received or paid.

Cancellation. This refers to the possible discarding of common components that are shared by the offered prospects. An example is the cancellation of a common probability-outcome pair, which is a restatement of Savage's sure-thing principle.

Segregation. An offered prospect with a riskless component such as a minimum gain (loss) is decomposed into a riskless component and the prospect with the riskless component taken from each outcome.

Combination. The probabilities associated with equal outcomes are combined to yield a single outcome with probability given by the sum of the respective probabilities.

Detection of Dominance. The dominated prospects are eliminated from the choice set prior to evaluation.

Simplification. This refers to the possible simplification of prospects by rounding off probability or outcome values.

The coding and the cancellation operations are prompted by the empirical evidence described in Chapter 3 section 4. As a result of coding, the outcome values of offered prospects are always stated in terms of gains and losses. Although more remains to be known about the conditions under which cancellation applies, incorporating such an operation does yield an additional degree of freedom in the description of choice phenomena. Even less is known about the simplification operation, which seems to be a plausible problem representation heuristic.

The other editing operations are related to the evaluation phase, which concerns the way a value function $v(x)$ of the outcome x and a decision weight function $\pi(p)$ of the probability of an outcome p combine to obtain the overall value of a prospect. The value function is concave for gains and convex for losses. The π function has the following properties:

- 1) π increases from $\pi(0)=0$ to $\pi(1)=1$.
- 2) $\pi(p) > p$, for small p .
- 3) $\pi(p) + \pi(1-p) < 1$, for $p \in (0, 1)$.
- 4) $\pi(pq)/\pi(p) \leq \pi(pqr)/\pi(pr)$, for $p, q, r \in (0, 1]$.

Prospect theory is developed for simple prospects of the form,

$$P \equiv p\delta_x + q\delta_y + (1-p-q)\delta_0, \quad (4.65)$$

which have at most two non-zero outcomes. If the outcomes are strictly positive (negative) and p and q add to unity, the simple prospect is known as a strictly positive (negative) prospect. A simple prospect is regular if it is neither strictly positive nor strictly negative. For regular prospects, the overall value V is obtained from the scales v and π in the following manner.

$$V(p\delta_x + q\delta_y + (1-p-q)\delta_0) = \pi(p)v(x) + \pi(q)v(y). \quad (4.66)$$

Unlike the von Neumann-Morgenstern utility, the value function $v(x)$ is a ratio scale, i.e. v vanishes at the reference point. This property was noted in Edwards (1961). There are however some difficulties with the use of the above expression, stemming from the nonlinearity of π . Since $\pi(p) + \pi(s-p)$ is not equal to $\pi(p') + \pi(s-p')$,

$$V(p\delta_x + (s-p)\delta_x + \delta_0) \neq V(p'\delta_x + (s-p')\delta_x + \delta_0). \quad (4.67)$$

In other words, two lotteries each yielding outcome x with probability s are not equivalent in preference. To get around this, the choice agent is assumed to apply the combination editing operation prior to evaluation.

The other difficulty necessitates the detection-of-dominance operation. Consider the following comparison:

$$\pi(p)v(x) + \pi(s-p)v(x+\epsilon) \quad ? \quad \pi(s)v(x) \quad (4.68)$$

Suppose the L.H.S. is less (more) than the R.H.S. for ϵ equal to zero, then the inequality would still hold for a small but positive (negative) ϵ . Thus, dominance is violated. Eliminating the dominated alternative before evaluation circumvents this problem, but not completely. Suppose

$$\pi(p)v(x) + \pi(s-p)v(x+\epsilon^0) < \pi(s)v(x), \quad (4.69)$$

for some positive ϵ^0 .

We can find $q^0 < s$ such that

$$\pi(p)v(x) + \pi(s-p)v(x+\epsilon^0) < \pi(q^0)v(x+\epsilon^0) < \pi(s)v(x). \quad (4.70)$$

This has the implication that

$$p\delta_x + (s-p)\delta_{x+\epsilon^0} + (1-s)\delta_0 \prec q^0\delta_{x+\epsilon^0} + (1-q^0)\delta_0,$$

$$\text{and} \quad ps\delta_x + (1-s)\delta_0 \succ q^0\delta_{x+\epsilon^0} + (1-q^0)\delta_0,$$

which is both normatively and empirically untenable.

For the strictly positive and strictly negative prospects, we will run into the same problem with violations of dominance if we adopt expression (4.66). Instead, Kahneman and Tversky proposed that people decompose a strictly positive (negative) prospect $P \equiv p\delta_x + (1-p)\delta_y$ with $y > x > 0$ ($0 > x > y$) into a riskless component δ_x and a risky component $P' \equiv p\delta_0 + (1-p)\delta_{y-x}$ and evaluate the segregated prospect via the following expression:

$$V(p\delta_x + (1-p)\delta_y) = v(x) + \pi(1-p)[v(y) - v(x)]. \quad (4.71)$$

The use of a different expression to evaluate strictly positive or strictly negative prospects leads to a new difficulty which is not covered by the editing operations. Consider the regular prospect $P \equiv p\delta_x + q\delta_{x+\epsilon} + (1-p-q)\delta_0$, with $x, \epsilon > 0$.

$$\begin{aligned} V(P) &= \pi(p)v(x) + \pi(q)v(x+\epsilon) \\ &< \pi(p)v(x) + \pi(1-p)v(x+\epsilon). \end{aligned}$$

Since V is concave for positive x ,

$$V(P) < \pi(p)v(x) + \pi(1-p)[v(x) + \epsilon v'(x)]. \quad (4.72)$$

Choose ϵ_0 such that

$$\epsilon_0 \leq \frac{v(x)}{v'(x)} \left[\frac{1 - \pi(p) - \pi(1-p)}{\pi(1-p)} \right] \quad (4.73)$$

This implies that,

$$p\delta_x + q\delta_{x+\epsilon_0} + (1-p-q)\delta_0 < \delta_x, \quad \forall q \in [0, 1-p).$$

For the case $q=1-p$, we apply expression (4.69) and obtain:

$$V(p) = v(x) + \pi(1-p)[v(x+\epsilon) - v(x)] > v(x). \quad (4.74)$$

This implies that

$$p\delta_x + (1-p)\delta_{x+\epsilon} > \delta.$$

The above implication is rather pathological. No matter how close q gets to $1-p$, the L.H.S. is strictly worse than δ_x . Yet, when q is

equal to $1-p$, the L.H.S. is strictly preferred to δ_x . We can get some feeling for the problem by estimating ϵ_0 for some reasonable values of x , p and $\pi(p)$. Suppose $x=\$100$, $p=.50$ and $\pi(0.50)=.45$. (Note that property (3) of the π function implies that $\pi(0.5)<0.5$). We have that,

$$\epsilon_0 \leq 10x \frac{[1-2\pi(0.5)]}{\pi(0.5)} < \frac{v(10)}{v'(10)} \frac{[1-2\pi(0.5)]}{\pi(0.5)}. \quad (4.75)$$

since the concavity of v implies that $x < \frac{v(x)}{v'(x)}$.

A conservative estimate for ϵ_0 is then given by

$$\epsilon_0 = 10 \times .1 \div .45 \approx 22.$$

Consider the following choice problems.

$$P_1 \equiv .5\delta_{\$100} + .45\delta_{\$122} + .05\delta_{\$0} \quad \text{vs.} \quad \delta_{\$100}$$

$$P_2 \equiv .5\delta_{\$100} + .49\delta_{\$122} + .01\delta_{\$0} \quad \text{vs.} \quad \delta_{\$100}$$

$$P_3 \equiv .5\delta_{\$100} + .499\delta_{\$122} + .001\delta_{\$0} \quad \text{vs.} \quad \delta_{\$100}$$

Prospect theory would predict that P_1 , P_2 , P_3 are all strictly worse than $\delta_{\$100}$. Moreover, the direction of preference remains unchanged as long as the probability of obtaining $\$122$ is less than 0.5! Note that the inequality, $x < \frac{v(x)}{v'(x)}$, used to obtain ϵ_0 is highly conservative since $v'(x)$ is a decreasing function for a concave $v(x)$. We can arrive at the above conclusion with a generally much higher ϵ_0 given a specific $v(x)$. If v is bounded, then $v'(x)$ tends to zero so that ϵ_0 can be made arbitrarily large if we consider a sufficiently large x .

Prospect theory builds on the form of the evaluation function first suggested by Edwards (1955). It treats systematically several classes of choice phenomena (cf. Chapter 3) that violate the implications of expected utility. However, the nonlinearity of the decision weight function π , which constitutes its main deviation from expected utility,

generates some serious difficulties for prospect theory. The immediate ones, namely the violation of the combination principle and stochastic dominance, were circumvented via the combination and the detection-of-dominance editing operations. Two problems however remain. One of the problems is the implication that there is always a prospect Q which is strictly worse than some prospect P but strictly better than another prospect P' that dominates P . The other problem has to do with the use of a different evaluation function for the strictly positive (negative) prospects. This gives rise to discontinuity in the preference represented which leads to untenable predictions of actual choice behavior.

4.5.3 Comparison

We conclude the section with a comparison of alpha utility with Allais' theory and prospect theory. The alpha utility representation $\Omega(F)$ can be stated in terms the expectation of a "value" function, v , with respect to an ' α -weighted' distribution F^α given by:

$$F^\alpha(x) = \int_{-\infty}^x \alpha dF / \int_{-\infty}^{\infty} \alpha dF,$$

where α is a strictly positive function. The role of α becomes clearer if we consider a simple distribution, $F \equiv \sum_{i=1}^n p_i \delta_{x_i}$. In this case,

$$\begin{aligned} \Omega(F) &= \int_{\mathbb{R}} v dF^\alpha \\ &= \sum_{i=1}^n q_i(F) v(x_i), \end{aligned}$$

$$\text{where } q_i(F) = p_i \alpha(x_i) / \sum_{j=1}^n p_j \alpha(x_j).$$

Like prospect theory, $\Omega(F)$ is obtained from the $v(x_i)$'s via a set of "decision weights", $\{q_i\}_{i=1}^n$, with each q_i being a nonlinear function of the probability p_i of obtaining the i th outcome, x_i . We should however note three distinctions. The q_i weights sum to unity but not the $\pi(p_i)$ weights of prospect theory. In addition to p_i , q_i depends on the rest of the p_j 's and all the x_j 's. Finally, the q_i -weights has the combination property, since if the j th and k th outcomes are both equal to x , then

$$\begin{aligned} &\Omega(p_j \delta_x + p_k \delta_x + \sum_{i \neq j,k} p_i \delta_{x_i}) \\ &= \frac{[p_j \alpha(x) + p_k \alpha(x)] v(x) + \sum_{i \neq j,k} p_i \alpha(x_i) v(x_i)}{\sum_{i=1}^n p_i \alpha(x_i)} \end{aligned}$$

$$\begin{aligned}
&= \frac{(p_j + p_k)\alpha(x)v(x) + \sum_{i=j}^n p_i \alpha(x_i)v(x_i)}{\sum_{i=1}^n p_i \alpha(x_i)} \\
&= \Omega((p_j + p_k)\delta_x + \sum_{i \neq j,k} p_i \delta_{x_i}).
\end{aligned}$$

Some of Allais' ideas can also be expressed in terms of the alpha utility representation. We can write $\Omega(F)$ as the sum of \bar{v} , the first moment of v with respect to F , and the deviation term $\int_R (v - \bar{v}) dF^\alpha$ shown below:

$$\Omega(F) = \bar{v} + \int_R (v - \bar{v}) dF^\alpha,$$

$$\text{where } \bar{v} = \int_R v dF.$$

Thus, the preference of an alpha utility decision maker depends on the first moment of v , and also the distribution of the deviation, $(v - \bar{v})$, of v from \bar{v} through the α -weighted distribution F^α . According to the proofs of the representation theorems in Chapters 1 and 2, the v -function, much like the von Neumann-Morgenstern utility, can be constructed from preference comparisons using a standard lottery. Allais may insist however that v be a psychological value function that is determined from introspective comparison among hypothetical changes in wealth position. In this case, we can apply the principle of Occam's razor to weed out the "psychological value" assumption since it is redundant.

Finally, we summarize via Table 4.2 some of the salient features of the theories treated in this chapter. A '+' sign means the corresponding theory is compatible with the property referred to. Otherwise, we use a '-' sign.

| | Properties | | | | Allais' Paradox |
|------------------|--------------|-----------|-------------|------------|--------------------|
| | Transitivity | Dominance | Combination | Continuity | |
| Expected Utility | + | + | + | + | - |
| Alpha Utility | + | + | + | + | + |
| Allais Theory | + | - | + | + | + |
| Prospect Theory | - | + | + | - | + |

Table 4.2: Comparison among Theories

A novelty of Prospect theory is the explicit use of editing operations prior to evaluation. Two of these operations are however needed to ensure consistency with dominance and combination. If Allais' theory were to adopt the detection-of-dominance editing operation, it would exhibit systematic intransitivity, as is the case for prospect theory (cf. expression (4.70)). Since lotteries are stated in terms of gains and losses, the coding operation seems a reasonable hypothesis about how people perceive monetary outcomes. Kahneman and Tversky (1979) have provided preliminary empirical evidence in support of two other editing operations, namely cancellation and segregation. These may be adopted by alpha utility theory if future empirical studies ascertain their validity as behavioral hypothesis.

CONCLUSION

CONCLUSION

5.1 SUMMARY

Part I of this dissertation contains the statements and proofs of two representation theorems. The first generalizes the quasilinear mean, M_ϕ , of Hardy, Littlewood and Polya:

$$M_\phi(F) = \phi^{-1}(\int_R \phi dF), \quad (5.1)$$

where ϕ is continuous and strictly monotone, and F is a probability distribution. We have weakened a characteristic property of the quasilinear mean, M_ϕ , called quasilinearity to obtain a more general mean value, $M_{\alpha\phi}$, which is characterized by an additional function α . The $M_{\alpha\phi}$ mean of a probability distribution F has the following form:

$$M_{\alpha\phi}(F) = \phi^{-1}(\int_R \alpha\phi dF / \int_R \alpha dF), \quad (5.2)$$

where α is continuous and strictly positive (negative). Through the strict inequality, $<$, binary relation, the $M_{\alpha\phi}$ mean induces a binary relation, \prec , among probability distributions:

$$F \prec G \iff M_{\alpha\phi}(F) < M_{\alpha\phi}(G). \quad (5.3)$$

The R.H.S. is equivalent to

$$\int_R \alpha\phi dF / \int_R \alpha dF < \int_R \alpha\phi dG / \int_R \alpha dG, \quad (5.4)$$

for a strictly increasing ϕ . Note that the ordering represented by (5.4) is more general than that represented by the expected utility representation, $\int_R \phi dF$. For convenience, we use $\Omega(F)$ to label the representation functional, $\int_R \alpha\phi dF / \int_R \alpha dF$.

An alternative approach to obtain the Ω representation, (5.4), is given in Chapter 2. Instead of probability distributions, we consider simple probability measures defined on some arbitrary set X . (A simple

probability measure is a convex linear combination of a finite number of point masses in X). Axioms were stated directly in terms of a binary relation, $<$, to obtain the corresponding Ω representation for simple probability measures:

$$\Omega(P) = E(\alpha v, P) / E(\alpha, P), \quad (5.5)$$

where P is a simple probability measure on a set X , and α and v are real-valued functions on X . This is contrasted with the approach based on axioms on mean values in Chapter 1. The reason why the mean value approach cannot be extended to simple probability measures on an arbitrary set X is that mean value itself may not be defined in X . As an example, consider the outcome set $X \equiv \{\text{status quo, being promoted, being fired}\}$. Being able to deal with an arbitrary outcome set such as the above example is an advantage for the simple probability measure approach. There are however certain drawbacks. Without further structural assumptions on X , we cannot discuss the notions of continuity and differentiability of the α and v functions nor can we generalize the Ω representation to include more general probability measures.

Part II of the dissertation concerns one specific area of application — decision theory. Interpreting the $M_{\alpha\phi}(F)$ mean of Chapter 1 as the certainty equivalent of a monetary lottery F , the corresponding induced binary relation, $<$, has the natural interpretation as 'strict preference' between lotteries. For non-monetary (finite) lotteries, we apply the representation theorem of Chapter 2. The hypothesis that the Ω representation of either approach represents a choice agent's preference among lotteries is referred to as alpha utility theory. This is logically equivalent to saying that the choice agent obeys either the mean value (certainty equivalent) axioms or the axioms on the ' $<$ ' (strict preference) binary relation.

Alpha utility theory is a generalization of expected utility theory in the sense that the expected utility representation is a special case of the alpha utility representation, and that alpha utility assumes a weaker form of a key property of expected utility, called substitutability (Pratt; Raiffa and Schlaifer, 1964), which is essentially the same as the quasilinearity property of Hardy, Littlewood and Polya. (A close counterpart of substitutability is the strong independence principle of Marschak (1950) and Samuelson (1952)),

The motivation for generalizing expected utility comes from difficulties it faced in the description of certain choice phenomena, especially the Allais paradox. These are summarized in Chapter 3.

Chapter 4 contains the formal statements of assumptions and the derivations of normative and descriptive implications of alpha utility theory. We stated conditions, taken from Chapter 1, for consistency with stochastic dominance and global risk aversion and derived a generalized Arrow-Pratt index of local risk aversion. We also demonstrated how a pair of α and v functions that satisfy both stochastic dominance and local risk aversion can be consistent with those choice phenomena, summarized in Chapter 3, that contradicts the implications of expected utility. The chapter ended with a comparison of alpha utility with two other theories that have attracted attention; namely, Allais' theory and prospect theory.

5.2 EXTENSIONS

We conclude by pointing out some potential areas of application. The quasilinear mean was given an interpretation as the 'equally-distributed-equivalent' level of income corresponding to an income distribution by Atkinson (1970) in his paper on the measurement of inequality. We may use the $M_{\alpha\beta}$ mean as a more general model of

equally-distributed-equivalent income. Is there any need for a more general measure? Consider two societies with income distributions F and G given by:

$$F \equiv 0.50\delta_{\$1,000} + 0.50\delta_{\$2,000}, \text{ and}$$

$$G \equiv 0.50\delta_{\$1,000} + 0.49\delta_{\$2,000} + 0.01\delta_{\$1,000,000}.$$

To many (including probably Mao Tse-Tung), society F fares better than society G. But the G distribution stochastically dominates the F distribution, so that $M_{\phi}(F)$ is less than $M_{\phi}(G)$. Therefore, the M_{ϕ} measure fails to reflect the relative welfare of the two societies for those who believe that society F is better off. This does not pose any difficulty for the $M_{\alpha\phi}$ measure since consistency with stochastic dominance is not an intrinsic property.

The departure of $M_{\alpha\phi}$ from M_{ϕ} can be made clearer if we consider a society of N individuals with incomes, $\{x_i\}_{i=1}^N$. The corresponding equally-distributed-equivalent is given by:

$$M_{\alpha\phi}(F) = \phi^{-1} \left(\frac{\sum_{i=1}^N \alpha(x_i) \vee(x_i)}{\sum_{i=1}^N \alpha(x_i)} \right). \quad (5.6)$$

A remarkable feature of (5.6) is the presence of complementarity across incomes of different individuals. That this is a desirable property is reinforced by the fact that an individual perceives concurrently the incomes of other individuals in the distribution whereas only one of a set of mutually exclusive outcomes will obtain in a lottery. We can think of the role of α as assigning discriminatory weights on individuals based on their attained incomes. Mao Tse-Tung's (hypothetical) preference for society F may then be explained in terms of a decreasing discrimination function α that treats wealthy individuals less 'equally' than the poorer folks.

In a recent paper on the measurement of poverty, Blackorby and Donaldson (1978b) applied Atkinson's equally-distributed-equivalent to the 'censored' distribution, i.e., the income distribution truncated at some exogenously established poverty line. This is tantamount to having a $M_{\alpha\phi}$ measure with an α that is constant up to the poverty line and zero beyond. It is natural to suggest that a decreasing α with an inflexion point at the poverty line (see Figure 5.1) would integrate the contribution due to the whole distribution and at the same time be particularly representative of the poorer folks with incomes below the poverty line.

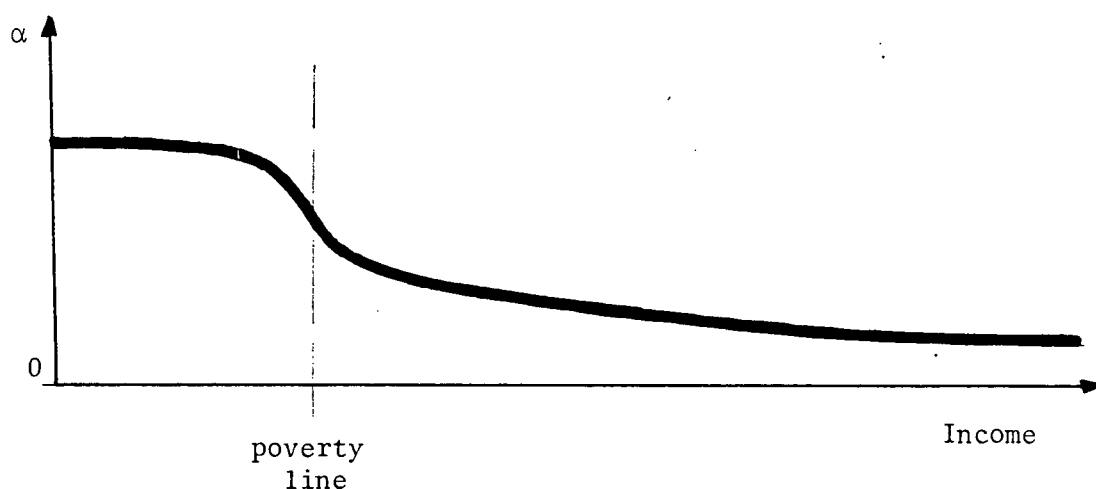


Fig. 5.1: An alpha function that discriminates against the rich

The $M_{\alpha\phi}$ mean can also be used to generate measures of inequality as follows:

$$\text{Relative Inequality Index} = 1 - M_{\alpha\phi}/\mu, \text{ and}$$

$$\text{Absolute Inequality Index} = \mu - M_{\alpha\phi},$$

where μ denotes the arithmetic mean. The relative index is due to Atkinson (1970) and the absolute index is due to Kolm (1976a,b). When the distribution of income is completely equal, both indices equal zero. The coefficient of variation, an occasional measure of income inequality, is related to $M_{\alpha\phi}$ by:

$$\text{coefficient of variation} = \{M_{1,1}(F) - \mu_F\}^{\frac{1}{2}}, \quad (5.7)$$

where $M_{s,t}(F) = \{\int_0^\infty x^{s+t} dF / \int_0^\infty x^s dF\}^{1/t}$ (i.e. $\alpha(x) = x^s$ and $\phi(x) = x^t$).

As an equally-distributed-equivalent, $M_{1,1}$ is undesirable because it is always higher than the arithmetic mean except at equality, and so encourages inequality. Also, its weighting function $\alpha(x) \equiv x$ assigns progressively more weight to the more wealthy.

The $M_{s,t}$ mean suggested above may be of use in statistics. We offer some examples.

$M_{2,1}$ = standard deviation \cdot coefficient of skewness, and

$M_{2,2}$ = standard deviation \cdot (coefficient of Kurtosis + 3) $^{\frac{1}{2}}$.

An equality that generalizes the well known result that the product of the arithmetic mean and harmonic mean is equal to the square of the geometric mean for two positive numbers, may be stated in terms of $M_{-t,2t}$ as follows.

$$M_{-t,2t}(F) = M_{\log x}(F), \quad (\text{geometric mean}) \quad (5.8)$$

where $F \equiv \sum \frac{1}{N} \delta_{x_i}$, with $x_i > 0$,

when the frequency polygon of $\log x_i$ is symmetrical about the axis of ordinate at $\frac{1}{N} \sum \log x_i$. Canning (1934) proposed the use of

$$\{M_{-t,2t}(F)/M_{\log x}(F)\}^{\frac{1}{2}}, \quad (5.9)$$

as a descriptive measure of asymmetry.

Harsanyi (1977) applied expected utility to an individual making a moral judgement about alternative social situations. Making a moral judgement, in this case, means making a hypothetical best choice under the assumption that the individual assumes the position of any one member of the society with equal chance. A social situation X , which is a listing of the N persons' states, would be perceived as a lottery that assigns the individual to any particular individual's position with $1/N$ chance. An expected utility decision maker would then maximize the expected utility corresponding to social situation X . This is simply the arithmetic average, $\sum_{i=1}^N u_i(X)/N$, of his von Neumann-Morgenstern utility, $u_i(X)$, for taking the position of the i th person. An alpha utility decision maker will however maximize a slightly more general expression:

$$\sum_{i=1}^N \alpha_i(X) v_i(X) / \sum_{i=1}^N \alpha_i(X), \quad (5.10)$$

where $\alpha_i(X)$ denotes his α value for being in i th person's shoes within the social situation X . This is a weighted average of the $v_i(X)$'s with weights given by the $\alpha_i(X)$'s. Expression (5.10) may be interpreted as the average of the $v_i(X)$'s but the alpha utility decision maker uses a 'biased' estimate, $\alpha_i(X) / \sum_{i=1}^N \alpha_i(X)$, for his probability of being the i th person. Again, the alpha utility expression has the advantage of allowing for complementarity, which is supported by our intuition about how moral judgements are made.

Several questions remain untouched. There are other situations where the representation functions have the additive structure of expected utility. One example is the averaging of individual utility

functions using a set of fixed weights to obtain a group utility function (Keeney and Raiffa, 1976). Another example is the time-honored practice of discounting a time stream of utilities (Koopmans, 1972). The intuition that complementarity should not be ruled out is strong in both cases. For these and other similar examples, the alpha utility representation provides a useful first candidate for a departure from the additive structure in order to incorporate complementarity across attributes.

The last decade saw a tremendous growth in the literature on the microeconomics of uncertainty. Practically all of these works are based on the assumption that choice agents maximize expected utility. We have not investigated what new results would follow or what old results are robust if we introduce alpha utility choice agents. One example is comparative risk aversion. The condition for one expected utility maximizer to always be more risk averse than another is given in Pratt (1964). An equivalent result appeared in Hardy, Littlewood and Polya (1934) in the form of an inequality between two quasilinear means. The corresponding result for alpha utility choice agents or the $M_{\alpha\phi}$ mean, however, remains an open question. Another example is Samuelson's (1967) conjecture, which was confirmed by Kalai and Schmeidler (1977), that Arrow's impossibility theorem will hold in a cardinal setup where individuals and society express their preferences by von Neumann-Morgenstern utility functions. It is perhaps safe to conjecture that Arrow's impossibility theorem would still hold if individuals and society express their preferences via alpha utility functions.

REFERENCES

- Aczel, J. (1966), Functional Equations and Their Application, Academic Press, New York and London.
- Allais, M. (1953), "Le comportement de l'homme rationnel devant le risque", Econometrica 21, No. 4, 503-546.
- Allais, M. (1979), "The So-Called Allais Paradox and rational decisions Under Certainty," in M. Allais and O. Hagen, eds., Expected Utility Hypothesis and the Allais Paradox, Reidel Publishing Company, 437-682.
- Allais, M. & O. Hagen (1979), Expected Utility Hypothesis & the Allais Paradox, D. Reidel Publishing Co.
- Anscombe, F.J. & R.J. Aumann (1963), "A Definition of Subjective Probability," Annals of Mathematical Statistics, Vol. 34, No. 1, 199-205.
- Arrow, K.J. (1971), Theory of Risk Bearing, Markham Publishing Company, Chicago.
- Atkinson, A.B. (1970) "On the Measurement of Inequality," Journal of Economic Theory, 2, 244-263.
- Becker, G.M. & C.G. McClintock (1967), "Value: Behavioral Decision Theory," Annual Review of Psychology, Vol. 18, 239-286.
- Bernoulli, D. (1954) "Specimen theoriae novae de mensura sortis," Comentarii academiae scientiarum imperialis Petropolitanae, Vol. 5, 175-192; translated as, "Exposition of a New Theory on the Measurement of Risk," Econometrica, Vol. 22, 23-26.
- Ben-tal, A. (1977) "On Generalized Means and Generalized Convex Functions," Journal of Optimization Theory and Applications, Vol. 21, No. 1, January, 1-13.
- Blackorby, C. and D. Donaldson (1978a), "A Theoretical Treatment of Ethical Indices of Absolute Inequality" forthcoming in International Economic Review, Department of Economics Discussion Paper 78-03, University of British Columbia.
- Blackorby, C. and D. Donaldson (1978b), "Ethical Indices for the Measurement of Poverty, Department of Economics Discussion Paper 78-04, University of British Columbia.
- Blackwell, D. (1951), "Comparison of Experiments," Proceedings of Symposium on Mathematical Statistics and Probability, 2nd, Berkeley, J. Neyman and L.M. Lecom, eds., University of California Press, 93-102.

- Blackwell, D. and M.A. Girshick (1954), Theory of Games and Statistical Decisions, Wiley & Sons, Inc.
- Canning, J.B. "A Theorem Concerning A Certain Family of Averages of a Certain Type of Frequency Distribution," Econometrica 2, 442. The abstract is a reporter's summary of an unpublished paper.
- Chew, S.H. (1979), "A Generalization of the Quasilinear Mean of Hardy, Littlewood and Polya," Institute of Applied Mathematics and Statistics Technical Report 79-39, University of British Columbia.
- Chew, S.H. and K.R. MacCrimmon (1979b), "Alpha Utility, Lottery Composition and the Allais Paradox," Faculty of Commerce and Business Administration Working Paper #686, University of British Columbia.
- Chew, S.H. and K.R. MacCrimmon (1979a), "Alpha Utility Theory: A Generalization of Expected Utility," Faculty of Commerce and Business Administration Working Paper #669, University of British Columbia.
- Dalton, H. (1920), "The Measurement of Inequality of Incomes," Economic Journal 20, 348-361.
- DeGroot, M.H. (1970), Optimal Statistical Decisions, McGraw-Hill Book Co., New York.
- Diamond, P. & M. Rothschild (1978) Uncertainty in Economics, Academic Press, New York.
- Edwards, W. (1954), "The Theory of Decision Making," Psychological Bulletin, Vol. 51, 380-417.
- Edwards, W. (1955), "The Prediction of Decisions Among Bets," Journal of Experimental Psychology 60, 265-277.
- Edwards, W. (1961), "Behavioral Decision Theory," Annual Review of Psychology 12, 473-498.
- Fishburn, P.C. (1970), Utility Theory for Decision Making, John Wiley, New York.
- Friedman, M. & L.J. Savage (1948), "The Utility Analysis of Choice Involving Risk," Journal of Political Economy, Vol. 56, 279-304.
- Frisch, R. (1926), "Suv une Probleme d'Economie Pure," Norsk Mathematisk Forenings Skrifter, 1.
- Hadar, J. and W.R. Russell (1969), "Rules for Ordering Uncertain Prospects," AER 49, 25-34.

- Hagen, O. (1979), "Towards a Positive Theory of Preferences Under Risk," in Expected Utility Hypothesis and the Allais Paradox, M. Allais and O. Hagen, eds., Reidel Publishing Company, 271-302.
- Handa, J. (1977), "Risk, Probabilities and a New Theory of Cardinal Utility," Journal of Political Economy, Vol. 85, No. 1, February, 97-122.
- Hanoch, G. and C. Levy (1969), "Efficiency Analysis of Choices Involving Risk," Review of Economic Studies 36, 335-346.
- Hardy, G.H., J.E. Littlewood & G. Polya (1934), Inequalities, Cambridge University Press.
- Harsanyi, J.C. (1977), Rational Behavior and Bargaining Equilibrium in Games and Social Situations, Cambridge: Cambridge University Press.
- Herstein, I.N. & J. Milnor (1953), "An Axiomatic Approach to Measurable Utility," Econometrica, Vol. 21(2), April.
- Howard, R.A. (1964), "The Foundations of Decision Analysis," IEEE Trans. SSC-4:211-19.
- Jensen, N.E. (1967), "An Introduction to Bernoullian Utility. I. Utility Functions," Swedish Journal of Economics.
- Kahneman, D. & A. Tversky (1979), "Prospect Theory: An Analysis of Decision Under Risk," Econometrica, March, 263-291.
- Kalai, E. and D. Schmeidler (1977), "Aggregation Procedure for Cardinal Preferences: A Formulation and Proof of Samuelson's Impossibility Conjecture," Econometrica 45, No. 6, September, 1431-1437.
- Karmarkar, U.S. (1978), "Subjectively Weighted Utility: A Descriptive Extension of Expected Utility Model," OBHP, Vol. 21, 61-72.
- Keeney, R.L. & H. Raiffa (1976), Decisions with Multiple Objectives: References and Value Tradeoffs, Wiley & Sons, Inc.
- Kolm, S.Ch. (1976a), "Unequal Inequalities I," Journal of Economic Theory 12, June, 416-442.
- Kolm, S.Ch. (1976b), "Unequal Inequalities II," Journal of Economic Theory 13, August, 82-111.
- Koopmans, T.C. (1972), "Representation of Preference Orderings Over Time," in Decision and Organization, C.B. McGuire and R. Radner, eds., North Holland Publishing Company, Amsterdam.
- Krantz, D.H., D.R. Luce & A. Tversky (1971), Foundations of Measurement Vol. 1, Academic Press, New York & London.

- Lehmann, E.L. (1955), "Ordered Families of Distributions," Annals of Mathematical Statistics 26, 399-419.
- Lichtenstein, S. & P. Slovic (1971), "Reversal of Preference Between Bids and Choices in Gambling Decisions," Journal of Experimental Psychology 89, 46-55.
- Luenberger, D.G. (1969), Optimization by Vector Space Methods, New York: John Wiley & Sons.
- MacCrimmon, K.R. (1965), "An Experimental Study of the Decision-Making Behavior of Business Executives," unpublished dissertation, University of California, Los Angeles.
- MacCrimmon, K.R. (1968), "Descriptive and Normative Implications of the Decision Theory Postulates," in K. Borch and J. Mossin, eds., Risk and Uncertainty, St. Martin's Press, New York, 3-32.
- MacCrimmon, K.R., John F. Bassler & William T. Stanbury (1972), Unpublished Results, Risk Study Project, University of British Columbia.
- MacCrimmon, K.R. & S. Larsson (1979), "Utility Theory: Axioms Versus 'Paradoxes'," in M. Allais and O. Hagen, eds., Expected Utility and the Allais Paradox, Holland: D. Riedel.
- Machina, M.J. (1980), "Expected Utility Analysis Without the Independence Axiom," unpublished working paper, Department of Economics, University of California - San Diego.
- Markowitz, H. (1952), "The Utility of Wealth," Journal of Political Economy, Vol. 60, 151-158.
- Marschak, J. (1950), "Rational Behavior, Uncertain Prospects, and Measureable Utility," Econometrica 18, 111-141.
- Marschak, J. & R. Radner (1972), Economic Theory of Teams, Yale University Press.
- Meginniss, J.R. (1977), "Alternatives to the Expected Utility Rule," unpublished Ph.D. dissertation, University of Chicago.
- von Neumann, J. & O. Morgenstern (1947), Theory of Games and Economic Behavior, Princeton University Press, 2nd Edition, Princeton.
- Norris, Nilan (1976), "General Means and Statistical Theory," The American Statistician, February, Vol. 30, No. 1.
- Pratt, J. (1964), "Risk Aversion in the Small and in the Large," Econometrica, Vol. 32, No. 1-2, January-April, 122-136.

- Pratt, J.W., H. Raiffa & R. Schlaifer (1964), "The Foundations of Decisions Under Uncertainty: An Elementary Exposition," JASA 59, 353-375.
- Preston, M.G. & P. Baratta, "An Experimental Study of the Auction Value of An Uncertain Outcome," Journal of Psychology 61, 183-193.
- Raiffa, H. (1968), Decision Analysis: Introductory Lectures on Choice Under Uncertainty, Reading, Massachusetts: Addison-Wesley.
- Raiffa, H. & R. Schlaifer (1961), Applied Statistical Decision Theory, Division of Research, Harvard Business School, Boston, 356.
- Ramsey, F.P. (1931), "Truth and Probability," (1926) in The Foundations of Mathematics, R.B. Braithwaite, ed., Humanities Press.
- Rothschild, M. & J.E. Stiglitz (1970), "Increasing Risk I. A Definition," Journal of Economic Theory 2, 225-243.
- Samuelson, P.A. (1952), "Probability, Utility, and the Independence Axiom," Econometrica 20, 670-78.
- Samuelson, P. (1967), "Arrow's Mathematical Politics," in Human Values and Economic Policy, S. Hook, ed., New York: New York University Press, 41-51.
- Savage, L.J. (1954), The Foundations of Statistics, Wiley, New York.
- Slovic, P., B. Fischhoff & S. Lichtenstein (1977), "Behavioral Decision Theory," Annual Review of Psychology, 39.
- Slovic, P. & A. Tversky (1974), "Who Accepts Savage's Axiom?" Behavioral Science 19, 368-373.
- Weber, R. (1980), Personal Communication.
- Weerahandi, S. and J. Zidik (1979), "A Characterization of the General Mean," forthcoming in the Canadian Journal of Statistics, Department of Mathematics Discussion Paper, University of British Columbia.
- Whitmore, G.A. (1970), "Third-Degree Stochastic Dominance," American Economic Review 60, 457-459.