ECONOMETRICA

VOLUME 50 MARCH, 1982 NUMBER 2

"EXPECTED UTILITY" ANALYSIS WITHOUT THE INDEPENDENCE AXIOM[1]

BY MARK J. MACHINA[2]

Experimental studies have shown that the key behavioral assumption of expected utility theory, the so-called "independence axiom," tends to be *systematically* violated in practice. Such findings would lead us to question the empirical relevance of the large body of literature on the behavior of economic agents under uncertainty which uses expected utility analysis. The first purpose of this paper is to demonstrate that the basic concepts, tools, and results of expected utility analysis do not depend on the independence axiom, but may be derived from the much weaker assumption of smoothness of preferences over alternative probability distributions. The second purpose of the paper is to show that this approach may be used to construct a simple model of preferences which ties together a wide body of observed behavior toward risk, including the Friedman-Savage and Markowitz observations, and both the Allais and St. Petersburg Paradoxes.

1. INTRODUCTION

AS AN APPROACH to the theory of individual behavior toward risk, the expected utility model is characterized by the simplicity and normative appeal of its axioms, the familiarity of the notions it employs (utility functions and mathematical expectation), the elegance of its characterizations of various types of behavior in terms of properties of the utility function (risk aversion by concavity, the degree of risk aversion by the Arrow-Pratt measure, etc.), and the large number of results it has produced. It is thus not surprising that most current theoretical research in the economics of uncertainty, as well as virtually all applied work in the field (e.g. optimal trade, investment, or search under uncertainty)[3] is undertaken in the expected utility framework.[4]

Nevertheless, the expected utility hypothesis is still a particular hypothesis concerning individual preferences over alternative probability distributions over wealth. In the years following its revival by von Neumann and Morgenstern in the *Theory of Games and Economic Behavior* **[99]**, it became generally recognized that expected utility theory depended crucially on the empirical validity of the

[1] An earlier version of this paper was presented to the 1979 North American Summer Meetings of the Econometric Society, Montreal, June, 1979.

[2] I would like to thank Vince Crawford, Peter Diamond, Mark Durst, Ted Groves, Frank Hahn, Klaus Heiss, Walt Heller, David Kreps, Andreu Mas-Colell, Eric Maskin, Jim Mirrlees, David Newbery, Mike Rothschild, anonymous referees, the Editor, and especially Franklin Fisher for helpful comments on this material. They are, of course, not responsible for errors. I am also grateful to the National Science Foundation and the Social Science Research Council for financial support.

[3] See, for example, Helpman and Razin [**42**] and Levhari and Srinivasan [**51**].

[4] The one significant exception to this statement is the "state preference" approach to behavior toward risk (see, for example, Debreu [**18**, Ch. 7] or Hirshleifer [**44**]). However, since this approach works with distributions of payoffs over states rather than with distributions of probability mass over payoffs, many of the issues discussed in the present paper do not bear directly on this approach.

so-called "independence axiom."[5] One of several equivalent versions of this axiom reads "a risky prospect A is weakly preferred (i.e. preferred or indifferent) to a risky prospect B if and only if a $p:(1-p)$ chance of A or C respectively is weakly preferred to a $p:(1-p)$ chance of B or C, for arbitrary positive probability p and risky prospects A, B, and C." In particular, the role of the other axioms of the theory, which essentially amount to the assumptions of completeness and continuity of preferences, is essentially to establish the *existence* of a continuous preference function over probability distributions, in much the same way as is done in standard consumer theory.[6] It is the independence axiom which gives the theory its empirical content by imposing a restriction on the *functional form* of the preference function. It implies that the preference function may be represented as the expectation with respect to the given distribution of a fixed utility function defined over the set of possible outcomes (i.e. ultimate wealth levels). In other words, the preference function is constrained to be a linear functional over the set of distribution functions, or, as commonly phrased, "linear in the probabilities."

The high normative appeal of the independence axiom has been widely (although not universally)[7] acknowledged. However, the evidence concerning its *descriptive* validity is not quite as favorable. The example of its systematic violation in practice which is perhaps best known to economists is the famous "Allais Paradox." This example (described below) consists of asking individuals to choose a most preferred prospect out of each of two specific pairs of risky prospects. Researchers have found that the particular choices made by the great majority of subjects in this situation violate the independence axiom, and hence are inconsistent with the hypothesis of expected utility maximization.

In addition, a large amount of research on the validity of the expected utility model has appeared in the psychology literature, where experimenters have similarly discovered that preferences are in general *not* linear in the probabilities. Edwards, in one of his reviews of this literature, asserted of expected utility maximization that "in 1954 it was already clear that it too [i.e. as well as expected value maximization] does not fit the facts" [**26**, p. 474].

Although these findings have led some researchers, both psychologists and economists, to propose alternative theories of behavior toward risk,[8] expected utility theory continues to be the dominant framework of analysis in the economics literature. Since it is likely to remain so in the future, it would seem crucial that we have some idea of the descriptive realism of the theory in light of the apparent invalidity of its key behavioral assumption. In other words, "*how*

[5] Although this axiom did not appear explicitly in the original von Neumann-Morgenstern axiom system, Malinvaud [**59**] has shown it to have been implicitly assumed in their pre-axiomatic formulation. Two important early formulations of the axiom are those of Marschak [**61**] and Samuelson [**80**], each of whom refer to similar work by other authors.

[6] See, for example, Debreu [**18**, Ch. 4].

[7] See the debate between Wold, Schackle, Savage, Manne, Charnes, and Samuelson on the *a priori* plausibility of the independence axiom in the October, 1952 issue of this journal, as well as the remarks in Allais [**3**, pp. 99–103].

[8] See, for example, the set of models discussed in Section 2.5 below.

robust are the concepts, tools, and results of expected utility theory to failures of the independence axiom?"

The first purpose of this paper is to demonstrate that expected utility analysis is in fact quite robust to failures of the independence axiom. Specifically, it is shown that, far from depending on the independence axiom (i.e. linearity of the preference functional), the basic concepts, tools, and results of expected utility analysis may be derived by merely assuming smoothness of preferences (i.e. that the preference functional is differentiable in the appropriate sense). This implies that while the independence axiom, and hence the expected utility hypothesis, may not be empirically valid, the implications and predictions of theoretical studies which use expected utility analysis typically *will* be valid, provided preferences are smooth. Several such results, including the Arrow-Pratt theorem, are formally proven for the general case of smooth preferences.

The second purpose of this paper is to demonstrate that this general analytic approach, termed "generalized expected utility analysis," may be used to construct a simple, yet evidently quite powerful model of individual behavior toward risk. Specifically, it is shown that two simple hypotheses concerning the shape of a fixed nonlinear preference functional over probability distributions serve to generate predictions consistent with (i) the typical behavior exhibited in the Allais Paradox, (ii) other experimental evidence regarding systematic violations of the independence axiom, (iii) the general observations on insurance and lotteries made by Friedman and Savage in their classic article on the expected utility hypothesis, (iv) the subsequent observation by Markowitz and others that preferences over alternative gambles are relatively independent of the level of current wealth (and hence that utility functions apparently shift when wealth changes), and (v) the typical behavior exhibited in the St. Petersburg Paradox and its generalizations. Thus, a number of seemingly unrelated aspects of behavior toward risk are seen to be jointly consistent with the hypothesis that the individual is maximizing a fixed preference functional defined over distributions, which in addition is particularly simple in shape.

Section 2 of this paper offers a historical overview of the expected utility model as a descriptive model, treating each of the above five behavioral observations, and discussing the various, and often ad hoc, modifications of the model which have been made to account for some of them. The applications of the tools and theorems of expected utility theory to the analysis of general nonlinear preference functionals is developed in Section 3. In Section 4 this approach is used to construct a simple model of preferences which is consistent with (and in some cases predicts) each of the above five aspects of behavior. Among other things, it is argued that this model offers (i) a simple characterization of the exact nature of observed violations of the independence axiom, (ii) a reconciliation of the relative independence of gambling behavior to current wealth with the hypothesis of a fixed preference ranking of probability distributions over ultimate wealth, and (iii) a resolution of the debate in the expected utility literature concerning the boundedness of the utility function. The paper concludes (Section 5) with some brief remarks on the topics of testing the model and applications of the analysis to the study of social welfare functionals.

2. EXPECTED UTILITY MAXIMIZATION AS A DESCRIPTIVE MODEL

In this section we consider several classes of observations concerning individual preferences over risky prospects, and give an account of how the expected utility model has been used, and in some cases adapted and modified, to account for these various types of behavior.

2.1. *Insurance, Lotteries, Skewness Preference, and the Friedman-Savage Hypothesis*

The primary motivation for the classic article by Friedman and Savage [**33**] came from their observations that "the empirical evidence for the willingness of persons of all income classes to buy insurance is extensive" [**33**, p. 285, or **91**, p. 66], that "the empirical evidence for the willingness of individuals to purchase lottery tickets, or engage in similar forms of gambling, is also extensive" [**33**, p. 286, or **91**, p. 67], and their belief that a large number of individuals purchase both.[9] They offer as a von Neumann-Morgenstern utility function which explains these particular observations one which has the form shown in Figure 1. The key aspect of such a utility function is that it is concave, and hence locally risk averse, about low outcome levels (i.e. low levels of ultimate wealth), linear (to a second order approximation) and hence locally risk neutral at the inflection point, and convex (locally risk loving) for high outcome values.[10]

In addition to its well known implications concerning the purchase of insurance and lottery tickets, another implication of the utility function of Figure 1, noted by Markowitz [**60**, p. 156], is that an individual with such a utility function will tend to prefer positively skewed distributions (ones with large right tails) over negatively skewed ones (ones with large left tails). The purchase of a lottery ticket, for example, induces a positively skewed distribution if initial wealth was certain, and insuring against a small probability-low outcome event transforms a negatively skewed distribution into a symmetric (certain) one. Since a mean preserving increase in risk (see [**74**]) which is "centered" in the upper tail of a symmetric distribution induces positive skewness, and one which is centered in the lower tail induces negative skewness, a preference for positive over negative skewness suggests that the individual will tend to prefer increases in risk in the upper tail of a given initial distribution of wealth over equivalent risk increases in the lower tail. Such a tendency is clearly an implication of the utility function of Figure 1.

The notion of a relative preference for (equivalently, a lower aversion to) risk increases in the upper rather than the lower tail of an initial distribution may be formalized by adopting the following definition:

[9] See also the comments of Adam Smith and Alfred Marshall in this regard quoted in [**33**, p. 284 or **91**, p. 65], as well as the reference to a distant relative of the author [**33**, p. 280 or **91**, p. 58].

[10] Mention should be made of the various attempts (e.g. Flemming [**32**], Hakansson [**39**], Kim [**47**], and Kwang [**50**]) to reconcile the simultaneous purchase of insurance and lottery tickets with the assumption of general risk aversion via such assumptions as indivisibility of expenditure, imperfect capital markets, etc.

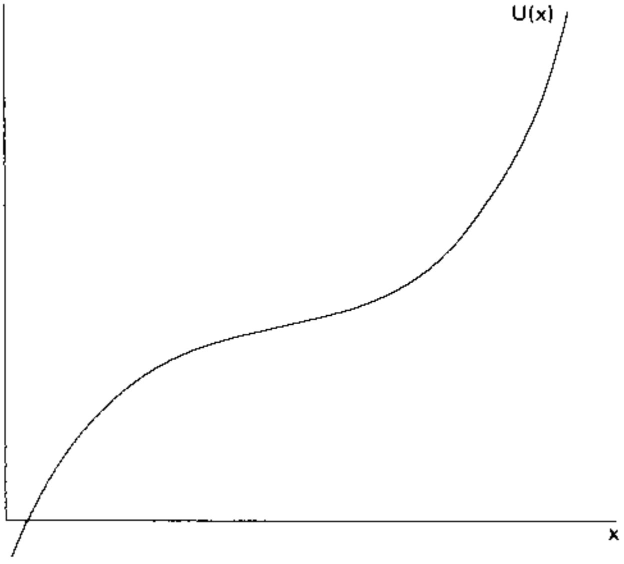

FIGURE 1

DEFINITION: If $F(\cdot)$ and $F^*(\cdot)$ are two cumulative distribution functions over a wealth interval $[0, M]$, then F^* is said to differ from F by a *simple compensated spread* if the individual is indifferent between F and F^*, and if $[0, M]$ may be partitioned into disjoint intervals I_L and I_R (with I_L to the left of I_R) such that $F^*(x) \geqq F(x)$ for all x in I_L and $F^*(x) \leqq F(x)$ for all x in I_R.[11]

A relative preference for risk increases in the upper rather than the lower tail of an initial distribution then implies that, if a given set of changes in the probabilities of the elements of the set $A \subset [0, M]$ can be represented as a sequence of simple compensated spreads, then the same respective changes in the probabilities of the set $A + c = \{x + c \mid x \in A\}$ are weakly preferred if the constant c is positive, and weakly not preferred if it is negative.[12]

[11] This definition is motivated by the "single crossing property" of Diamond and Stiglitz [**19**], and it is clear that when the individual is an expected utility maximizer, sequences of simple compensated spreads are equivalent to mean utility preserving increases in risk [**19**, pp. 341–345].

[12] It is important to distinguish between this behavioral principle and the Kahneman and Tversky "reflection effect" [**46**, pp. 268–269], which states that the preference ranking over a pair of prospects (defined in terms of gains and losses) reverses when all the outcome values are reversed in sign. Since such an effect concerns the relative rankings within two *distinct* pairs of prospects, and since any spread of probability mass relating the initial pair of prospects is itself "reflected," it is quite distinct from the present principle, which concerns the ranking of a *single* pair of prospects, each of which is obtained from a given initial distribution by a spread which, though horizontally translated, is not reflected. Note that while Kahneman and Tversky's associated hypothesis of "risk aversion in the positive domain [i.e. among prospects involving gains] . . . accompanied by risk seeking in the negative domain" [**46**, p. 268] is supported by their examples 7, 7′, 10, 11, 12, 13, and 13′, preferences in problems 1, 3, and 3′ may be explained by positive skewness preference, and in problems 2, 4, and 4′ by the differences in the expected values of the prospects. Examples 8, 8′, 14, and 14′, on the other hand, actually contradict their hypothesis.

There is evidence to suggest that positive skewness preference and a relative preference for risk increases in the upper rather than the lower tails of distributions are also exhibited by an important class of individuals not characterized by the utility function of Figure 1, namely global risk averters. Tsaing [**94**, pp. 359–360] and Hirshleifer [**45**, pp. 282–283] have argued that positive skewness preference is evidently prevalent among risk averse investors, the former pointing to a number of financial devices which allow investors to increase the positive skewness of their returns. Indeed, such preferences were espoused as long ago as the eighteenth century by Condorcet (see [**82**, pp. 44–45]). Evidence of a relative preference for risk increases in the upper as opposed to the lower tail of an initial distribution has also been uncovered by Mosteller and Nogee. At one point in their experiment [**66**, pp. 386–389], subjects were asked to leave written instructions to an "agent" who would be faced with a sequence of gambling opportunities in their absence. Although these instructions were predominantly risk averse, they frequently suggested that the agent play more liberally when doing well. In other words, there were some gambles the agent was instructed always to take, and some, never to take. Such a policy would result in some particular distribution of winnings. The designation of additional gambles which should be taken only if cumulative winnings have been high enough indicates that there are some further increases in risk which would be preferred if they occurred in the upper tail of this distribution, but not preferred if they occurred in the lower tail.[13]

2.2. *The St. Petersburg Paradox, the Structure of Lotteries, and the Boundedness of Utility*

At a later point in their article [**33**, pp. 296–297, or **91**, pp. 84–85], Friedman and Savage point out that an individual with a utility function as in Figure 1 and with initial wealth near the inflection point would always pay more for a lottery ticket offering a probability p of $\$Z$ than for a ticket with two such chances (i.e. probability $2p$) of winning $\$Z/2$. On this basis, they reject the shape in Figure 1 as inconsistent with their final observation, namely that (lottery designers are presumably profit maximizers, and) "lotteries typically have more than one prize" [**33**, p. 294, or **91**, p. 80]. Writers from Cournot (see [**90**, n. 127]) through Menger [**63**, p. 226] and Markowitz [**60**, pp. 153–154] have made essentially this same point, namely that the amount the individual would pay for a $1/n$ chance of winning $\$nZ$, though possibly increasing at first, is an eventually declining function of n. In light of this, Friedman and Savage modified their original proposed shape so as to include a terminal concave section, as in Figure 2.[14,15]

[13] Markowitz [**60**, pp. 155–156] has noted that such instructions also imply what has been seen to be a related behavior, namely positive skewness preference.

[14] Strictly speaking, the terminal segment must be *sufficiently* concave (see [**33**, n. 34]).

[15] Markowitz [**60**] subsequently modified the theory further by adding a third inflection point to the left of the first one, since "the individual generally will prefer one chance in ten of owing \$10,000,000 rather than owing \$1,000,000 for sure" [**60**, p. 154]. Thus, the amount the individual would pay to avoid a $1/n$ chance of losing $\$nZ$ may similarly eventually decline in n. An alternative explanation is that the individual views the actual consequences of owing either amount as identical (i.e. total bankruptcy) and simply acts to minimize the probability of this common outcome.

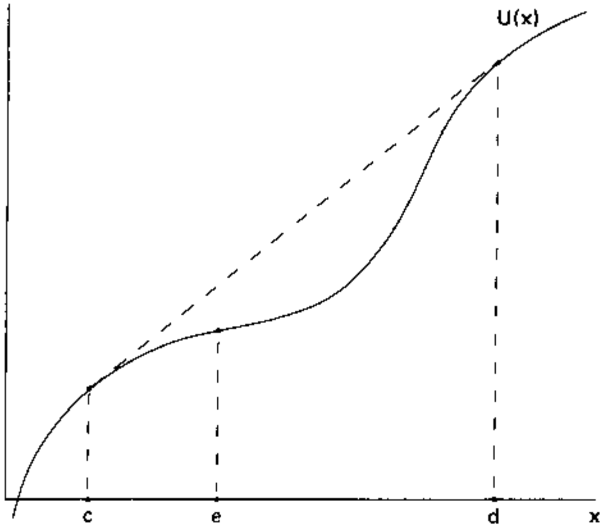

FIGURE 2

A second objection to the utility function of Figure 1 comes from the typical response to the famous "St. Petersburg Paradox" and its generalizations.[16] The original form of this paradox was the observation that an individual typically would never forgo a significant amount of wealth to engage in the gamble which offered a payoff of $\$2^{i-1}$ with probability 2^{-i} for $i = 1, 2, \ldots,$ even though the expected winnings from this gamble are infinite. Since an individual with a Figure 1 utility function clearly *would* forgo any finite sure level of wealth to take this gamble, such a utility function must be abandoned as unrealistic. In his classic article, Menger [**63**] generalized the paradox by showing that whenever the utility function was unbounded, similar gambles could be constructed which also had infinite certainty equivalents,[17] so that the utility function of Figure 2 must be further modified so as to be bounded for all outcome levels. More recently, Arrow [**6**, pp. 63–69] (see also Samuelson [**82**, pp. 35–36 and footnote 14]) has shown that an individual with unbounded utility must violate either the completeness or the continuity axiom of expected utility theory.[18]

A common objection to the "evidence" posed by the St. Petersburg Paradox and to the extent of the problems posed by unbounded utility has been that no person, or for that matter, no society, could ever offer such a gamble to the individual, and therefore it is meaningless to ask how much such a gamble would be worth. However, as has been shown (see Aumann [**9**, p. 444] and Samuelson [**82**, pp. 32–34]), the incompatibility of unbounded utility with "reasonable"

[16]See Samuelson [**82**] for a historical and critical overview of the literature surrounding the paradox, and Shapely [**86**, **87**] and Aumann [**9**] for more recent comments.

[17]Let $x_i = U^{-1}(2^i)$ and consider the gamble which offers $\$x_i$ with probability 2^{-i} for $i = 1, 2, 3, \ldots$.

[18]See also Ryan [**79**], Arrow [**7**], Shapley [**86**, **87**], Aumann [**9**], Fishburn [**29**], and Russell and Seo [**78**] on this issue.

behavior may be demonstrated even if only distributions with finite numbers of outcomes are considered. The simplest such instance is the implication that, if utility is unbounded, for any arbitrarily large amount $\$C$ and arbitrarily small positive probability p, there will always be some amount $\$Z$ such that the individual will prefer a p chance of winning $\$Z$ to a certain gain of $\$C$.

The evidence thus suggests that the utility function of Figure 1 must be replaced by one as in Figure 2, that $U(\cdot)$ must be bounded, and furthermore that the second inflection point must occur at an empirically relevant outcome level.[19] Although such restrictions are necessary to make the expected utility model consistent with the observations considered above, they reduce the elegance with which the observations of Section 2.1 were modelled by the utility function of Figure 1. In particular, the degree of risk aversion is no longer monotonic in the outcome level. Thus, for example, a Mosteller-Nogee subject with a Figure 2 utility function would instruct an agent to play more liberally when doing well, *provided winnings have not been too high*, and, if playing conservatively at this high wealth level results in sufficient losses, more liberal gambles ought once again to be taken.

2.3. *The Relative Invariance of Gambling Behavior to Initial Wealth and the Markowitz Hypothesis*

The next objection to (and modification of) the original Friedman-Savage utility function concerned not so much the typical shape of the utility function, but rather the more fundamental issue of the stability of preferences. Recall that the independence axiom, in conjunction with the other axioms of expected utility theory (see, for example, Herstein and Milnor [**43**]) implies that the preference ranking corresponds to the expectation of a fixed utility function defined over final consequences, or in other words, ultimate levels of wealth. Indeed, Friedman and Savage, in their discussion of the standard method of estimating the utility function by fixing its values at two arbitrary wealth levels, pointed out that the expected utility hypothesis would be violated if the use of another pair of wealth levels as reference points "yielded a utility function differing in more than origin and unit of measure from the one initially obtained" [**33**, p. 292, or **91**, pp. 77–78]. Thus, when faced with alternative gambles, that is, prospects expressed in terms of *deviations* from current wealth, the individual will choose that gamble whose implied distribution over ultimate wealth levels has the highest expected utility.[20] This procedure of "integrating" (i.e. convoluting) alternative gambles with initial wealth before ranking is referred to by Kahneman and Tversky as "asset integration" [**46**, p. 264].

[19] Stiglitz [**92**] has argued that the requirement of boundedness does not rule out the case of $U(x)$ being convex for all x less than a trillion dollars. If such were the case, however, we would not observe lotteries offering multiple prizes of values less than this amount, nor would the individual's valuation of a $1/n$ chance of $\$n$ start declining until n were at least one trillion.

[20] Hence Edwards' statement that "the fundamental idea of a utility scale is such that the whole structure of a subject's choices [over such gambles] should be altered as a result of [the change in initial wealth due to] each previous choice (if the choices are real ones involving money gains or losses)" [**24**, p. 395].

However, as noted by Markowitz [**60**], the assumption that the utility function of Figure 2 is defined over ultimate wealth levels is *not* consistent with the observed tendency of individuals of all wealth levels to purchase insurance and lottery tickets.[21] Individuals with wealth levels less than c ("poor") or greater than d ("well to do") would never accept any fair bets, for example, yet "even poor people, apparently as much as others, buy sweepstakes tickets, play the horses, and participate in other forms of gambling. Rich people play roulette and the stock market" [**60**, p. 153]. Similarly, an individual with wealth just below d would be willing to take an expected loss for the privilege of *underwriting* insurance against large losses. In addition, individuals with wealth near $(c + d)/2$ would prefer all symmetric and other fair bets of up to at least $(d - c)/2$, even though "generally people avoid symmetric bets" [**60**, p. 154]. Noting that individuals of all wealth levels tend to behave as if their initial wealth was near the left inflection point e in Figure 2, Markowitz hypothesized that changes in wealth caused the utility function to shift horizontally so as to keep this inflection point at or near the current or "customary" level of wealth.[22]

The experimental evidence similarly suggests that individual gambling behavior at different initial wealth levels is more indicative of a shifting utility function than of movements along a fixed utility function. In reestimating the "utility curves" of subjects after periods of a few days to several weeks (during which their wealth must surely have changed by amounts greater than those involved in the experiment), Davidson, Suppes, and Siegel found that seven of their eight subjects "gave responses which were substantially consistent with the original results" and that three of them "performed the rather astonishing feat of exactly duplicating their first choices (they were given no hint as to what their earlier choices had been)" [**17**, pp. 68–69, 81]. Since Mosteller and Nogee also failed to account for wealth changes between sessions, their conclusion that "on the basis of empirical curves [constructed from data obtained over several sessions] it is possible to estimate future behavior in comparable but more complicated risk-taking situations" [**66**, p. 403] also supports this conclusion.[23] In a somewhat different context, Edwards [**23**] observed preferences over pairs of prospects involving fixed probabilities and a common (though variable) expected value and noted that "if the utility curve is non-linear . . . then a markedly different set of choices should be made at each different EV-level (since at each different EV-level different amounts of money, falling at different places on the utility curve, are involved in the bets)" [**23**, p. 87]. Finding that the observed choices generally did not depend on the expected value level, he was led to reject the existence of "one utility curve consistent with all these sets of choices" [**23**, p. 87].

[21] This implication was also noted by Friedman and Savage [**33**, pp. 300–301 or **91**, pp. 90–91] (see also Hirshleifer [**44**, pp. 259–261]).

[22] Markowitz suggested that the utility function might also undergo a horizontal expansion as it shifts to the right, so that the distance between the inflection points might be an increasing function of initial wealth [**60**, p. 155].

[23] Note that neither Davidson, Suppes, and Siegel nor Mosteller and Nogee found that individuals typically exhibited constant absolute risk aversion, which would also have served to explain their observations.

Presumably as a result of their survey data, Kahneman and Tversky have also concluded that "the preference order of prospects [defined in terms of gains and losses] is not greatly altered by small or even moderate variations in asset position" [**46**, p. 277]. Most recently, Binswanger [**12**] has used experimentally obtained data on the risk preferences of rural Indian villagers to conduct an explicit test of the asset integration hypothesis, which was formally rejected in favor of the alternative of a shifting utility function.[24]

The Markowitz hypothesis of a shifting utility function implies that changes in initial wealth essentially cause the individual to go back and *rerank* the entire "consumption set" of distributions over ultimate wealth levels. Such a hypothesis, asserting that preferences cannot be defined independently of the current consumption point is, in the words of Eden, "disturbing to economists who use the assumption of 'constant tastes' quite heavily . . . it is hard to see how positive economics can do without this assumption and it is almost impossible to think of welfare economics without it" [**20**, p. 125]. While the phenomenon of a relative invariance of gambling behavior to initial wealth, and in particular a simultaneous propensity to insure, buy lottery tickets, and avoid symmetric bets at all wealth levels may well contradict the joint hypothesis of constant tastes *and expected utility maximization*, such behavior (including the insurance-lotteries-symmetric bets observation) is not incompatible with the existence of *any* fixed preference ranking over ultimate wealth distributions, as will be shown in Sections 4.4 and 4.5 below. Thus, before dropping the assumption of constant tastes in order to save the assumption that the individual is maximizing the expectation of some utility function at each initial wealth level, it is crucial that we examine the extent to which this latter assumption is in fact warranted by the data.[25]

[24]The evidence on the effect of changes in wealth *within* sessions, however, is less conclusive. In an analysis of some of the subjects of their pilot study, Mosteller and Nogee found at least some evidence that the greater the amount of money "on hand," the greater the propensity to gamble [**66**, pp. 399–402], although that portion of the evidence which they present seems inconclusive, and Edwards has in fact interpreted them as concluding that "the amount of money possessed by the subjects did not seriously influence their choices" [**24**, pp. 395]. Mosteller and Nogee's analysis of the original Preston and Baratta data, on the other hand, "did *not* reveal . . . any evidence of differential bidding for gambles at the beginning and end of the game [i.e. session]" [**66**, p. 398]. Similarly, while McGlothlin found a tendency for bettors at pari-mutuel horse races to increase both the size of their wagers and the proportion of long-shot bets during the course of the racing day (i.e. "session"), he also found that, with the exception of the seventh ("feature") race of the day and the final eighth race (where "bettors apparently refrain from making bets which would not recoup their losses if successful" [**62**, p. 614]), "the first six races all yield E-vs.-odds patterns that do not differ from the pattern for the total sample by more than the sampling error" [**62**, p. 610]. Since intra-session wealth changes are due solely to gambling gains and losses, differences in the short and long run effects of such changes might be related to Davidson, Suppes, and Siegel's observations that "winning or losing several times in a row made subjects sanguine or pessimistic and tended to produce altered responses to the same offers" and "if the same syllable [on a random die] came up three times in succession, for example, the subjective probability would temporarily decrease for most subjects" [**17**, pp. 53, 54].

[25]See Section 4.4, however, for references to some experimentally observed choice behavior (under both certainty and uncertainty) of a different nature which apparently *does* contradict the assumption of constant tastes.

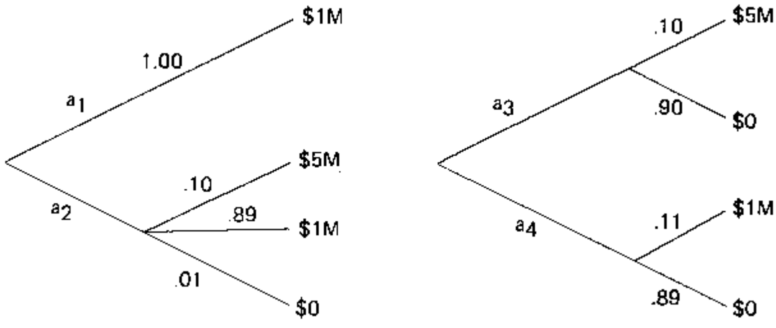

FIGURE 3.—The Allais Paradox ($1M = $1,000,000).

2.4. *Systematic Violation of the Independence Axiom: The Allais Paradox*

In this section and the next we consider the evidence that, even at a fixed initial asset position, individual rankings over alternative risky prospects tend to *systematically* violate the independence axiom, and hence are inconsistent with the hypothesis of expected utility maximization.

The most widely discussed of such examples is the famous "Allais Paradox" (see, for example, Allais [**2**, **3**, **4**], Allais and Hagen [**5**], Raiffa [**70**, pp. 80–86], or Morrison [**64**]), where the individual is asked to rank a particular pair of risky prospects a_1 and a_2, and then asked to rank the pair a_3 and a_4, where the payoffs and their corresponding probabilities are given in Figure 3. Since the shifts in probability mass in moving from prospect a_1 to a_2 and from a_4 to a_3 both consist of lowering the probability of winning $\$1M$ by .11 and raising the probabilities of winning $\$5M$ and $0 by .10 and .01 respectively, an expected utility maximizer would either prefer a_2 to a_1 and a_3 to a_4 (i.e. prefer the common shift) if the sign of $[.01\,U(w) - .11\,U(w+1M) + .10\,U(w+5M)]$ is positive, or else prefer a_1 to a_2 and a_4 to a_3 (i.e. not prefer the shift) if the sign is negative, where w is initial wealth.

Allais and others (e.g. Raiffa [**70**, p. 80], Morrison [**64**]), however, have found that the majority of subjects questioned prefer a_1 in the first pair and a_3 in the second, a pair of choices which violates the independence axiom. Morrison, for example, reported that when presented to a class of first year MBA students who had not been exposed to expected utility theory, 80 per cent made the above choices, and that even when presented to a similar class which had been exposed to the theory, the percentage of such "inconsistent" choices was still 50 per cent. Indeed, Savage himself made these choices when presented with the example, although he later changed his preferences to conform with the independence axiom [**83**, pp. 101–103].[26] The fact that the *same* pair of choices are made by so high a percentage of subjects makes the Allais Paradox a key example of the systematic violation of the independence axiom. Finally, it should be noted that

[26] Note that the version of the paradox presented in [**83**] differs from Figure 3 in that the labeling of prospects 3 and 4 is reversed and all payoffs are scaled down by $\frac{1}{2}$.

this example is not an isolated case: individuals faced with similar choice situations have tended to violate the axiom in what will be shown to be the same systematic fashion (see, for example, the evidence reported in Kahneman and Tversky [**46**, Problems 1 & 2 and Table 1], Hagen [**38**, pp. 285–296], and MacCrimmon and Larsson [**57**, pp. 350–369], most of which involves more moderate payoff levels than in the Allais Paradox).

One characterization of how such behavior systematically violates the independence axiom involves comparing the class of utility functions which rank a_1 over a_2 with the class of functions which rank a_3 over a_4. Note that the prospects a_1 and a_2 respectively stochastically dominate[27] a_4 and a_3, and recall that a utility function $U(\cdot)$ ranks a_1 over a_2 (a_3 over a_4) if and only if $[.01U(w) - .11U(w + 1M) + .10U(w + 5M)]$ is negative (positive), or equivalently, if and only if receiving $\$1M$ with certainty is preferred (not preferred) to a 10/11 chance of $\$5M$. Thus, in evaluating the change from a_1 to a_2, the typical individual acts as if using a utility function which is more risk averse than the one "used" to evaluate the change from a_4 to a_3. An analysis of the above cited evidence of Kahneman and Tversky, Hagen, and MacCrimmon and Larsson similarly reveals a tendency for individuals to violate the independence axiom by ranking the stochastically dominating pair of prospects "according to" a utility function which is more risk averse than the one "used" to rank the stochastically dominated pair.[28]

An alternative characterization of such behavior, in a form more directly comparable to the independence axiom, involves the notion of the "conditional certainty equivalent" of a prospect. Returning to Figure 3, define the prospect a^* as a 1/11:10/11 chance of winning \$0 or $\$5M$ respectively, and let E be an event with probability .11. Then the prospects a_1, a_2, a_3, and a_4 have the same distributions as the compound prospects which respectively yield $\$1M$, a^*, a^*, and $\$1M$ if E occurs, and $\$1M$, $\$1M$, \$0, and \$0 if $\sim E$ occurs. It is clear that the independence axiom requires that the conditional certainty equivalent of a^* in E, that is, the amount which the individual would, *ex ante*, just be willing to substitute for a^* if E occurs, be independent of what would ensue if $\sim E$ were to occur. However, the typical preference for a_1 over a_2 and a_3 over a_4 implies that the conditional certainty equivalent of a^* in E is less than $\$1M$ when $\sim E$ yields $\$1M$ with certainty and greater than $\$1M$ when $\sim E$ yields \$0. A similar analysis of Kahneman and Tversky [**46**, Problems 1 & 2] and MacCrimmon and Larsson [**57**, pp. 360–369] (i.e. that portion of the above cited evidence which can be formulated in this framework) also reveals the general property that, for a given event E and prospect a^*, stochastically dominating shifts in the conditional distribution of wealth in $\sim E$ will lower the conditional certainty equivalent of a^*

[27] Throughout this paper, "stochastic dominance" refers to first order stochastic dominance (see Hadar and Russell [37]).

[28] Note that in some of these examples the vectors of changes in the probabilities of the payoffs between each pair are not identical (as in the Allais Paradox) but rather scalar multiples of each other, a fact which has no bearing on the applicability of the above type of calculation.

in E. Thus, contrary to the precepts of the independence axiom, the more that individuals stand to lose if the event E occurs (that is, the better off they would be in $\sim E$), the more risk averse they become in evaluating a given risky prospect a^* in E. Equivalently, individuals are less risk averse toward a given prospect a^* in E if E is the "preferred event" (i.e. when $\sim E$ involves low outcome values) than when E is not the preferred event (i.e. when $\sim E$ involves high outcome values).[29]

A possible objection to the validity of this (and the following) evidence against the independence axiom is that individuals, when shown how their choices violated the axiom, would, like Savage, change their preferences to conform with it (see the discussions in Savage [**83**, pp. 102–103], Raiffa [**70**, pp. 80–86], and MacCrimmon [**56**, pp. 9–11]). While this phenomenon would clearly be a testimony to the *normative* appeal of the axiom, it is irrelevant to the positive theory of behavior toward risk (would an insurance company base its estimate of the pedestrian fatality rate on the widely held belief that the individual, *if reminded*, would always choose to look both ways before crossing a street?). Finally, there is evidence that the ability of experimenters to talk subjects out of preferences which violate the independence axiom may not be due to its "intuitive appeal" so much as the subject's desire to conform with the explicit or implicit beliefs of the experimenter. MacCrimmon [**56**, pp. 9–11] and Slovic and Tversky [**88**] reported that, when presented with opposing written arguments, subjects whose initial choices conformed to the axiom were about as likely to change their preferences as subjects who initially violated it.[30]

2.5. *Systematic Violation of the Independence Axiom: Oversensitivity to Changes in Small Probabilities and the Subjective Expected Utility Hypothesis*

The third important characterization of how the independence axiom is systematically violated, namely that, relative to expected utility maximization, individuals are oversensitive to changes in the probabilities of small probability-outlying events, may also be illustrated by the Allais Paradox. Note that the

[29] It is important to distinguish this type of behavior from that discussed in Section 2.1. Roughly speaking, the current aspect is that the individual's aversion to the riskiness of a^* in E grows with a general rise in the payoff levels in $\sim E$, whereas the earlier aspect was that it drops if there is a uniform rise in the payoffs in E (i.e. a uniform rise in the payoff levels of a^* itself).

[30] Although in a similar study Moskowitz found that presenting subjects with opposing written arguments and allowing them to discuss these among themselves led to a net decrease in the proportion of violations of the axiom, nevertheless 73 percent of the initial "Allais type" preference rankings expressed by subjects remained unchanged after the discussions [**65**, pp. 232–237, Table 6]. (When the written arguments were presented but no discussion was allowed, he found no net change in the degree of conformity with the axiom and a "persistency rate" of Allais type choices of 93 percent [**65**, p. 234, Tables 4 & 6].) Moskowitz also found that, of the three alternative forms of representing the choice problem he presented, that form which was judged the "clearest representation" by the majority of subjects (the "tree" diagram) led to the lowest degree of conformity with the axiom, the highest proportion of Allais type violations, and the highest persistency rate of these violations [**65**, pp. 234, 237–238].

common shift from a_1 to a_2 and from a_4 to a_3 may be thought of as moving .10 units of probability mass from the outcome $w + 1M$ to the outcome $w + 5M$ and moving .01 units of mass from $w + 1M$ to w. When the initial prospect is a_1, the upward movement of the .10 mass is not enough to compensate for the downward movement of the .01 mass, and the shift is not preferred. However, when the initial prospect is a_4, the outcome w is no longer such an "outlying event" of the initial distribution, since (relative to a_1) its probability has increased from 0 to .89. As a result, the individual is no longer as sensitive to the .01 rise in the probability of this event (at the expense of the preferred event $w + 1M$) and this downward movement of mass is now more than compensated by the upward movement of the .10 mass, so the shift (to a_3) is preferred.

Alternatively (and as will be seen below, equivalently), changing the initial prospect from a_1 to a_4 may be viewed as making the outcome $w + 5M$ "more outlying" relative to w and $w + 1M$, since, although the probability of this outcome hasn't changed, in moving from a_1 to a_4 a probability mass of .89 has moved *farther away* from the outcome level $w + 5M$. Thus, with the outcome $w + 5M$ more of an outlying event in the distribution a_4 than in a_1, the individual is now more sensitive to changes in its probability, and the upward movement of mass from $w + 1M$ to $w + 5M$ is now more than enough to compensate for the downward movement from $w + 1M$ to w, so the shift becomes preferred. A similar analysis of the evidence of Kahneman and Tversky, Hagen, and MacCrimmon and Larsson cited in the previous section also reveals this general tendency for individuals to be "oversensitive" to changes in the probabilities of low probability-outlying events.

A second source of evidence that individuals violate the independence axiom via a systematic oversensitivity to the probabilities of low-probability events are the empirical fittings by both psychologists and economists of the so-called "subjective expected utility" models.[31] Such models assume that the individual transforms the known set of objective probabilities $\{p_i\}$ of a risky prospect into their corresponding "subjective probabilities" $\{\pi(p_i)\}$ (called "decision weights" by Kahneman and Tversky [**46**]) and then maximizes the value of $\sum_i x_i \cdot \pi(p_i)$ ("subjective expected value" or SEV) or the value of $\sum_i U(x_i) \cdot \pi(p_i)$ ("subjective expected utility" or SEU), where p_i is the probability of the outcome value x_i. Since the independence axiom requires that $\pi(p_i)$ be linear, empirical estimates of the $\pi(p_i)$ function would yield information regarding the nature of any systematic violation of the axiom.

Such studies have on the whole found that, relative to linearity, individuals overemphasize small probabilities and underemphasize large probabilities. Applications of the SEV model to a wide range of both experimentally and nonexperimentally generated data have consistently yielded estimated $\pi(p)$ functions which are proportionately greater from small values of p than for large ones (see

[31] A systematic presentation and discussion of this class of models is given in Edwards [**25**, **27**] (see also Wallsten [**100**] and the references cited there, as well as the surveys of Edwards [**24**, **26**] and Luce and Suppes [**54**]). Modified versions of these models have recently been introduced into the economics literature by Handa [**40**] (see also Fishburn [**30**]) and Kahneman and Tversky [**46**].

for example Preston and Baratta [**69**], Griffith [**36**], Sprowls [**89**], Nogee and Lieberman [**67**], and Ali [**1**]). Although Ali [**1**] and others have argued that an estimated $\pi(p)$ function which overweights small probabilities is exactly what we would expect if the SEV model (which constrains the outcome values x_i to enter in linearly) were (mis)applied to choice data generated by an expected utility maximizer with terminally convex utility, Edwards has shown in another context that observed nonlinear "probability preferences" cannot be completely accounted for by utility considerations alone (Edwards [**21**; **22**, p. 66; **23**, pp. 84–95; **25**, pp. 211–212]). Experiments by Edwards [**25**] and Tversky [**95**, **96**] designed to overcome this problem by obtaining joint estimates of $\pi(p_i)$ and $U(x_i)$ in the SEU model continued to reveal a preponderant tendency towards overemphasizing small probabilities relative to larger ones.[32] Finally, in a somewhat different type of experiment designed to distinguish between behavior due to the curvature of the utility function and that due to exaggeration of small probabilities, Yaari [**101**] found that "acceptance sets" for bets were generally convex, which ruled out the possibility of convexities in the utility function, and implied that the risk loving behavior exhibited by seven of his seventeen subjects can only be explained (in the SEU framework, at least) by an exaggeration of the small probabilities of the favorable outcomes in these gambles. Although Rosett [**71**, **72**] has subsequently argued that the experimental design in [**101**] was not sufficient to rule out the existence of convex portions of the utility function, he noted that his objection did not apply to Yaari's conclusion regarding the exaggeration of small probabilities [**71**, p. 535; **72**, pp. 77–82], and indeed has also obtained evidence of such exaggeration in a subsequent experiment of his own [**73**, pp. 489, 492].

Since $\pi(0)$ must necessarily equal zero, a tendency for individuals to deviate from a linear $\pi(p)$ function in the direction of a relative overemphasis of small probabilities implies that, at least for values of p below a certain level, $\pi(p)$ must be a concave function of p. Since the sensitivity to a change in the probability of an outcome value x_i in the SEU model is given by $U(x_i) \cdot \pi'(p_i)$, this evidence reaffirms the principle that the individual is more sensitive to changes in the probabilities of events when their initial probabilities are low than when they are high.[33]

Although the SEU model allows for a relatively straightforward estimation of the individual's relative sensitivity to changes in low versus high probabilities, it

[32] In other experimental applications of the SEU model, Wallsten obtained mixed evidence on whether $\pi(p)$ differed from p by more than a scale factor [**100**, p. 39] and, though they conducted no formal estimation, Lichtenstein [**52**, p. 168] and Kahneman and Tversky [**46**, p. 281] similarly concluded that individuals overweight small probabilities.

[33] Some researchers (e.g. Preston and Baratta [**69**, p. 188]) have found that the slope of $\pi(p)$ may start rising again for values of p near unity. This would reflect the fact that, as the probability of the outcome value x_i approaches one, the probabilities of all other outcomes must go to zero, and as a result, the individual becomes increasingly sensitive to shifts which increase the probability of x_i at the expense of these other outcome values. In other words, the effect of a given shift of probability mass from x_j to x_i (which equals $U(x_i)\pi'(p_i) - U(x_j)\pi'(p_j)$) is large in magnitude when either $p_i \simeq 1$ and $p_j \simeq 0$ or when $p_i \simeq 0$ and $p_j \simeq 1$.

exhibits many undesirable properties. Once $\pi(p)$ is nonlinear, for example, behavior is no longer characterized by the shape of $U(\cdot)$ alone, and the main results of expected utility theory (such as the characterization of risk aversion by the concavity of $U(\cdot)$) no longer apply. More important, however, is the fact that, except in the case when it reduces to expected utility, the SEU model is incapable of incorporating the property of monotonicity (i.e. a preference for stochastically dominating distributions) in the sense that any individual maximizing $\sum_i U(x_i) \cdot \pi(p_i)$ with a nonlinear $\pi(p)$ function will *necessarily* prefer some distributions to ones which stochastically dominate them.[34] Similarly, unless $\pi(p)$ is linear, no subjective expected utility maximizer can exhibit general risk aversion (i.e. aversion to all mean preserving increases in risk), even over restricted ranges of possible outcomes.[35] In the author's view, this *intrinsic* incompatibility of the SEU model with the plausible behavioral properties of risk aversion, and especially general monotonicity, makes it unacceptable as a descriptive model of behavior toward risk.

It is useful to keep in mind the distinction between an oversensitivity to changes in the probabilities of small probability events and any tendency, under conditions of uncertainty rather than risk, to *overestimate* the probabilities of rare events. Since in this section and the preceding one we have treated behavior in situations where the individuals are told the relevant probabilities, this latter tendency, while it may exist, is irrelevant to the behavior considered here. Similarly, note that the principle of oversensitivity to changes in the probabilities of small probability-outlying events is not contradicted by the fact that individuals often tend to neglect altogether (i.e. treat as impossible) events of very low probability (see the references cited in Arrow [**6**, p. 14] and Samuelson [**82**, pp. 39–40]). The neglect (for all practical purposes) on an increase in the probability of disaster from 0 to .0000001 would only violate this principle if the same absolute increase in the probability of disaster was *not* neglected when the initial probability was .5000000.[36]

[34] As a result of their proof of this, Kahneman and Tversky [**46**, pp. 283–284] modify their model to require that the stochastically dominated distributions be eliminated from the choice set before the rest of the alternatives are ranked by their modified SEU function. However, they point out that this process permits what they call "indirect violations of dominance" ([**46**, p. 284]) and may result in intransitive choices.

[35] To see this, note that a mean preserving spread of probability mass from the outcome x_2 to the outcomes $x_1 = x_2 - t$ and $x_3 = x_2 + t$ will not be preferred if and only if $[U(x_1) \cdot \pi'(p_1) - 2U(x_2) \cdot \pi'(p_2) + U(x_3) \cdot \pi'(p_3)]$ is nonpositive, which will be true for all p_1, p_2, p_3, and small t if and only if $\pi'(p)$ is constant and $U(\cdot)$ is concave. It is straightforward to verify that this incompatibility with general risk aversion (as well as with general monotonicity) extends to all "additive" models with a maximand of the form $\sum_i f(x_i, p_i)$ where $f(\cdot, \cdot)$ is smooth and not identically equal to $U(x_i) \cdot p_i$ for some $U(\cdot)$.

[36] Note finally that the violations of expected utility discussed in this section and the preceding one cannot be explained by merely observing that individual rankings are often stochastic. Such "random" preferences over risky prospects were noted by Mosteller and Nogee [**66**] and have been explicitly incorporated into the expected utility model by Fishburn [**28**, **31**] (see also Luce and Raiffa [**53**, pp. 371–384] and the references mentioned there). While randomness clearly characterizes real life choice, stochastic *expected utility* models cannot account for the systematic violations of the independence axiom which have been considered, since such models would predict that, in the Allais Paradox for example, either a_1 and a_4 are chosen most of the time, or else a_2 and a_3 are.

3. THE ANALYSIS OF GENERAL NONLINEAR PREFERENCE FUNCTIONALS

In this section we demonstrate the robustness of expected utility analysis to violations of the independence axiom by showing how the fundamental concepts, tools, and results of expected utility theory may be applied to the general case of an individual possessing a "smooth" preference ranking over alternative probability distributions over ultimate wealth.

3.1. *Smooth Preferences and the "Local Utility Function"*

We take as our choice set the set $D[0, M]$ of all probability distribution functions $F(\cdot)$ over the interval $[0, M]$ and assume that the individual's preference ranking over this set is complete, transitive, and representable by a real-valued preference functional $V(\cdot)$ on $D[0, M]$.[37] Throughout this paper, all integrals will be taken over the interval $[0, M]$ unless otherwise specified.

For the purpose of defining continuity of preferences, the most appropriate topology to place on $D[0, M]$ is the topology of weak convergence, which defines a sequence $\{F_n(\cdot)\} \subset D[0, M]$ as converging to $F(\cdot)$ if and only if $F_n(x) \to F(x)$ at each continuity point x of $F(\cdot)$.[38] This topology renders as convergent the following sequences, each of which economic agents are likely to "think of" as convergent: (i) pointwise convergence of the density functions of a sequence of continuous distributions, (ii) the "collapse" of a sequence of distributions to the degenerate distribution $G_c(\cdot)$, which from now on will be used to denote the distribution which assigns unit mass to the point c, and (iii) the convergence of the sequence $\{G_{c_n}(\cdot)\}$ to $G_c(\cdot)$, where $c_n \to c$. Finally, since it may be shown that a sequence $\{F_n(\cdot)\}$ converges to the distribution $F(\cdot)$ in this topology if and only if $\int g(x)dF_n(x) \to \int g(x)dF(x)$ for all continuous $g(\cdot)$ on $[0, M]$, the weak convergence topology is the weakest (i.e. coarsest) topology on $D[0, M]$ for which the expected utility functional $\int U(x)dF(x)$ is continuous for all continuous $U(\cdot)$ on $[0, M]$.

The condition of differentiability of $V(\cdot)$ requires in addition the existence of a norm on the space $\Delta D[0, M] = \{\lambda(F^* - F) \mid F, F^* \in D[0, M], \lambda \in R^1\}$. Lemma 1 in the Appendix shows that the weak convergence topology on $D[0, M]$ is in fact induced by the L^1 metric $d(F, F^*) \equiv \int |F^*(x) - F(x)| dx$, which induces the norm $\|\lambda(F^* - F)\| \equiv |\lambda| \cdot d(F, F^*)$ on $\Delta D[0, M]$.[39]

Adopting this norm, our differentiability or "smoothness" condition will be that the preference functional $V(\cdot)$ be *Fréchet differentiable on the space* $D[0, M]$ *with respect to the norm* $\|\cdot\|$. Fréchet differentiability is the natural notion of differentiability on spaces such as $D[0, M]$ (i.e. subsets of Banach spaces),[40] and the function $V(\cdot)$ is said to be Fréchet differentiable at the point F in $D[0, M]$ if

[37] We assume throughout this section that the outcome space $[0, M]$ is bounded. In particular, note that the metric we shall define on $D[0, M]$ is only applicable if this is the case.

[38] See, for example, Billingsley [**10**, **11**].

[39] This follows since $\Delta D[0, M]$ is a linear subspace of $L^1[0, M]$ and $\|\cdot\|$ is just the L^1 norm restricted to this subspace.

[40] See, for example, Rudin [**77**, p. 248] or Luenberger [**55**, pp. 172–177].

there exists a continuous linear functional $\psi(\cdot\,; F)$ defined on $\Delta D[0, M]$ such that

$$\lim_{\|F^* - F\| \to 0} \frac{|V(F^*) - V(F) - \psi(F^* - F; F)|}{\|F^* - F\|} = 0. \tag{1}$$

In particular note that convergence here is required to be uniform in $\|F^* - F\|$.[41]

An equivalent method of representing this notion is to write

$$V(F^*) - V(F) = \psi(F^* - F; F) + o(\|F^* - F\|), \tag{2}$$

where $o(\cdot)$ denotes a function which is zero at zero and of a higher order than its argument. By footnote 39 and the Hahn-Banach theorem, there exists a continuous linear extension of $\psi(\cdot\,; F)$ to $L^1[0, M]$. Thus, by the Riesz representation theorem on $L^1[0, M]$,[42] we have that for any $F^* \in D[0, M]$,

$$\begin{aligned} \psi(F^* - F; F) &= \int (F^*(x) - F(x)) h(x; F)\, dx \\ &= -\int (F^*(x) - F(x))\, dU(x; F), \end{aligned} \tag{3}$$

where $h(\cdot\,; F) \in L^\infty[0, M]$ and

$$U(x; F) \equiv -\int_0^x h(s; F)\, ds, \tag{4}$$

from which it follows that $U(\cdot\,; F)$ is absolutely continuous and hence differentiable almost everywhere on $[0, M]$ (see Klambauer [**48**, p. 122]).

Substituting (3) into (2) and integrating by parts (see Lemma 2 in the Appendix) yields

$$V(F^*) - V(F) = \int U(x; F)(dF^*(x) - dF(x)) + o(\|F^* - F\|). \tag{5}$$

From (5) we see that a differential movement from the distribution $F(\cdot)$ to a distribution $F^*(\cdot)$ changes the value of the preference functional $V(\cdot)$ by $\int U(x; F)(dF^*(x) - dF(x))$, that is, by the difference in the expected value of $U(x; F)$ with respect to the distributions $F^*(\cdot)$ and $F(\cdot)$. In other words, in ranking differential shifts from an initial distribution $F(\cdot)$, the individual acts *precisely as would an expected utility maximizer*, with "local utility function" $U(x; F)$.[43] Intuitively, the fact that any Fréchet differentiable preference function

[41] Note that this is a stronger requirement than just that the directional derivative exist for all directions $F^* - F$ and be linear in the direction. This latter condition, known as Gateaux differentiability (see Luenberger [**55**, pp. 171–172]), is not even sufficient to ensure continuity.

[42] See, for example, Klambauer [**48**, p. 172] or Royden [**76**, p. 103].

[43] Note that the local utility function at a distribution $F(\cdot)$ displays the usual affine invariance properties of a von Neumann-Morgenstern utility function, since from (5) it is clear that neither an additive nor a multiplicative transformation of $U(\cdot\,; F)$ will alter the ranking of differential shifts from $F(\cdot)$. Note that by analogy with standard indifference curve analysis, however, the local utility functions $U(\cdot\,; F)$ and $U(\cdot\,; F^*)$ of the indifferent distributions $F(\cdot)$ and $F^*(\cdot)$ can only be used to compare respective differential shifts from these distributions if the functions $U(\cdot\,; F)$ and $U(\cdot\,; F^*)$ are not subjected to different multiplicative transformations.

may be thought of as "locally expected utility maximizing" follows from the fact that differentiable functions are "locally linear," and that for preference functionals over probability distributions, linearity is equivalent to expected utility maximization.[44]

The simplest example of such a nonlinear preference functional is the specification

$$\tilde{V}(F) \equiv \int R(x)\,dF(x) + \tfrac{1}{2}\Big[\int S(x)\,dF(x)\Big]^2 \tag{6}$$

$$= E_F\big[R(x)\big] + \tfrac{1}{2}\big[E_F\big[S(x)\big]\big]^2,$$

which may be termed "quadratic in the probabilities,"[45] and with local utility function

$$\tilde{U}(x;F) = R(x) + S(x)\Big[\int S(z)\,dF(z)\Big] = R(x) + S(x)E_F\big[S(z)\big], \tag{7}$$

where $E_F[\cdot]$ denotes expectation with respect to the probability distribution $F(\cdot)$.[46] Thus, an individual with such a preference function would prefer a differential shift from the distribution $F(\cdot)$ to a distribution $F^*(\cdot)$ if and only if the sign of $[E_{F^*}[\tilde{U}(x;F)] - E_F[\tilde{U}(x;F)]]$ is positive.

3.2. *The Mathematical Characterization of Behavior*

While the function $U(\cdot;F)$ may be used to rank differential shifts from an initial distribution $F(\cdot)$, in general there will be no neighborhood of $F(\cdot)$ in $D[0,M]$, however small, over which the ranking induced by the local utility function corresponds exactly to the ranking induced by the preference functional itself. Nevertheless, the present extension of expected utility analysis may similarly be applied to nondifferential (i.e. global) situations in much the same manner in which standard multivariate calculus may be used to show that a nonlinear but differentiable function will exhibit certain global properties (such as monotonicity) throughout a region provided its linear approximations at every point in the region exhibit the property in question, even though the linear approximations at different points in the region will in general be different linear functions. In other words, in a large body of cases, if the appropriate qualitative property (e.g. concavity) holds for every local utility function throughout a region, then the preference functional will display the corresponding behavioral

[44] An earlier special case of this result, proven in Samuelson [**81**, pp. 34–37] and discovered by the author in the course of writing this paper, is that an individual with "smooth" preferences will rank alternative differential deviations of the payoff levels from an initially certain distribution according to expected *value* maximization. This follows from the present result coupled with the fact that expected utility maximizers with differentiable utility functions will rank such differential changes in the payoffs according to expected value.

[45] This functional form can be shown to be a special case of the most general quadratic form $\frac{1}{2}\int\int T(x,z)\,dF(x)\,dF(z)$ where without loss of generality we may assume $T(x,z) \equiv T(z,x)$, and with local utility function $\int T(x,z)\,dF(z)$.

[46] We assume $R(\cdot)$ and $S(\cdot)$ to be absolutely continuous with $R'(\cdot), S'(\cdot) \in L^\infty[0,M]$.

property (e.g. risk aversion) throughout the region, *even though the local utility functions are not the same throughout the region* (i.e. even though the individual is not an expected utility maximizer).

The general method by which such results can be proven is the use of path integrals in the space $D[0,M]$. Specifically, if the path $\{F(\cdot;\alpha)\,|\,\alpha\in[0,1]\}$ is smooth enough so that the term $\|F(\cdot;\alpha)-F(\cdot;\alpha^*)\|$ is differentiable in α at $\alpha=\alpha^*$, then from equation (5) we have

$$\text{(8)}\qquad \frac{d}{d\alpha}(V(F(\cdot;\alpha)))\Big|_{\alpha^*}=\frac{d}{d\alpha}\left(\int U(x;F(\cdot;\alpha^*))\,dF(x;\alpha)\right)\Big|_{\alpha^*}+\frac{d}{d\alpha}(o(\|F(\cdot;\alpha)-F(\cdot;\alpha^*)\|))\Big|_{\alpha^*}=\frac{d}{d\alpha}\left(\int U(x;F(\cdot;\alpha^*))\,dF(x;\alpha)\right)\Big|_{\alpha^*},$$

since the derivative of the higher order term $o(\cdot)$ will be zero at zero. Combining (8) and the Fundamental Theorem of Integral Calculus yields that

$$\text{(9)}\qquad V(F(\cdot;1))-V(F(\cdot;0))=\int_0^1\left[\frac{d}{d\alpha}\left(\int U(x;F(\cdot;\alpha^*))\,dF(x;\alpha)\right)\Big|_{\alpha^*}\right]d\alpha^*,$$

which illustrates how the individual's reaction to the shift from $F(\cdot;0)$ to $F(\cdot;1)$ will depend on the characteristics of the local utility function at each point (i.e. distribution) along the path $\{F(\cdot;\alpha)\,|\,\alpha\in[0,1]\}$. As a first application of this method, we have the following theorem.

THEOREM 1: *Let $V(\cdot)$ be a Fréchet differentiable preference function on $D[0,M]$. Then $V(F^*)\geqq V(F)$ whenever $F^*(\cdot)$ stochastically dominates $F(\cdot)$ if and only if $U(x;F)$ is nondecreasing in x for all $F(\cdot)\in D[0,M]$. (Proof in Appendix.)*

To ensure strict monotonicity (i.e. strict preference for stochastic dominance) we shall assume from now on that $U(x;F)$ is strictly increasing in x for all $F(\cdot)$ in $D[0,M]$. This would be true in the quadratic example of equations (6) and (7), for example, if $R(x)$ was strictly increasing and $S(x)$ either nonnegative and nondecreasing or else nonpositive and nonincreasing. Consider now Theorem 2.

THEOREM 2: *Let $V(\cdot)$ be a Fréchet differentiable preference function on $D[0,M]$. Then $V(F^*)\leqq V(F)$ whenever $F^*(\cdot)$ differs from $F(\cdot)$ by a mean preserving increase in risk if and only if $U(x;F)$ is a concave function of x for all $F(\cdot)\in D[0,M]$. (Proof in Appendix.)*

Thus, a sufficient condition for the quadratic preference functional $\tilde{V}(\cdot)$ of equation (6) to exhibit global risk aversion is that $R(x)$ be concave and $S(x)$ either everywhere concave and nonnegative or else everywhere convex and nonpositive.

Theorem 2 has two important implications for the generality of expected utility theory, which follow from the "if" and "only if" parts of the theorem, respectively. The first is that researchers, who in order to study the behavior of risk averters in various situations have modelled them as expected utility maximizers with concave utility functions, are likely to have proven results which are also valid in the more general case of smooth preferences. The second is that concavity of a cardinal function of wealth is a *complete characterization* of risk aversion, in the sense that *any* risk averter must possess concave local utility functions, whether or not he or she is an expected utility maximizer. Thus, the researcher who would like to drop the expected utility hypothesis and study the nature of general risk aversion can apparently work completely within the framework of expected utility analysis.

3.3. *Behavioral Equivalencies*

Besides its elegant characterizations of types of behavior in terms of mathematical properties of the utility function, another of the useful aspects of expected utility theory is the behavioral equivalencies it implies. Indeed, it is *only* those theorems which relate various types of behavior which are ultimately meaningful, and the only reason one studies the behavior implied by, say, a concave utility function in some situation is because of the behavior it implies or to which it is equivalent in other situations.

It is in this respect, however, that the independence axiom would seem to be instrumental in deriving results in expected utility theory. For the independence axiom is essentially a global restriction on preferences, as it implies that the local utility functions at all distributions $F(\cdot)$ in $D[0, M]$ are identical. Thus, for example, knowing that an individual is averse to small mean preserving spreads about all certain (i.e. degenerate) distributions implies that the common local utility function is concave, which by Theorem 2 implies that the individual is averse to increases in risk about all initial distributions. Clearly, however, such a result no longer holds when the independence axiom is replaced by the *local* assumption of smoothness of preferences.

Nevertheless, it is possible to prove various behavioral equivalencies in the general case of smooth preferences and, as with Theorems 1 and 2, although these results do not require the independence axiom, they do follow the basic structures of the corresponding expected utility results. As a first example, consider again the expected utility result that aversion to all mean preserving increases in risk is implied by the local condition of aversion to all mean preserving spreads about certain (degenerate) distributions and the global restriction imposed by the independence axiom, which requires that if $F^*(\cdot)$ is weakly preferred to $F(\cdot)$, then the distribution $(1-p)F^{**}(\cdot)+pF^*(\cdot)$ will be weakly preferred to $(1-p)F^{**}(\cdot)+pF(\cdot)$, for arbitrary p, F, F^*, and F^{**}. Together these conditions imply (but are *not* implied by) the condition that for arbitrary p, F, and F^{**}, the distribution $(1-p)F^{**}(\cdot)+pG_{\mu_F}(\cdot)$ is weakly preferred to $(1-p)F^{**}(\cdot)+pF(\cdot)$ (where μ_F is the mean of F), i.e. in any compound lottery, the individual would always prefer to substitute (ex ante) the mean of any of the possible risky prizes for the risky prize itself. Note that this last condition requires merely that the conditional certainty equivalent (see Section 2.4) of the distribution F always be no greater than its mean, and not that it necessarily be some constant value independent of p and F^{**}, as does the independence axiom. The following theorem shows that in the case of behavioral equivalencies as well, the expected utility result provides the complete structure of the corresponding more general result, but that the sort of global equality condition imposed by the independence axiom may be replaced by the weaker requirement that the appropriate qualitative condition (i.e condition (i) of the theorem) hold throughout.

THEOREM 3: *The following properties of a Fréchet differentiable preference function $V(\cdot)$ on $D[0, M]$ are equivalent: (i) for arbitrary distributions $F(\cdot), F^{**}(\cdot) \in D[0, M]$ and arbitrary probability p, $V((1-p)F^{**} + pG_{\mu_F}) \geqq V((1-p)F^{**} + pF)$, where μ_F is the mean of F; (ii) $U(x; F)$ is concave in x for all $F \in D[0, M]$; and (iii) if $F^*(\cdot)$ differs from $F(\cdot)$ by a mean preserving increase in risk, then $V(F^*) \leqq V(F)$. (Proof in Appendix.)*

As a final example, we consider the Arrow-Pratt theorem of expected utility theory, which, as extended by Diamond and Stiglitz [**19**, Theorem 3], relates the mathematical condition of different levels of the Arrow-Pratt measure of absolute risk aversion to the behavioral conditions of differing certainty equivalents of risky prospects, effects of compensated increases in risk, and demands for a risky asset. Once again, the independence axiom (i.e. the requirement of constant conditional certainty equivalents) may be replaced by the requirement that the conditional certainty equivalents of one individual are always no greater than the corresponding conditional certainty equivalents of the other individual, regardless of whether the conditional certainty equivalents of either individual are constant (i.e. independent of p and F^{**}).

Proceeding similarly, we define the "conditional demand for a risky asset" as the value of α which yields the most preferred distribution in the set $\{(1-p)F^{**} + pF_{(1-\alpha)r+\alpha\tilde{z}} \mid \alpha \in R^1\}$,[47] where r is a positive constant and $\tilde{z}$ a nonnegative random variable with mean greater than r, i.e. as the optimal proportion of a portfolio to place in the risky asset when there is some probability $(1-p)$ that for exogenous reasons (such as bankruptcy) the distribution of wealth will be $F^{**}(\cdot)$ regardless of the composition of the portfolio. While the independence axiom requires that such conditional demands be a constant independent of p or F^{**}, we require merely that the conditional demands for one individual always be no greater than the corresponding ones of the other individual, regardless of whether these conditional demands vary with p or F^{**}.

[47] Where $F_{(1-\alpha)r+\alpha\tilde{z}}(\cdot)$ stands for the distribution function of $(1-\alpha)r + \alpha\tilde{z}$, etc.

In order to confine our study of asset demand to the case of "regular" optima, we adopt the following condition—a generalization of the condition that indifference curves in the (σ, μ) plane be upward sloping and bowed downward—which serves to rule out both risk lovers and "plungers" as in the classic study of Tobin [**93**, pp. 77–78].

DEFINITION: A risk averse individual is said to be a *diversifier* if, for all distributions $F^{**}(\cdot)$, positive probabilities p, positive constants r, and nondegenerate nonnegative random variables $\tilde{z}$, the individual's preferences over the set of distributions $\{(1-p)F^{**} + pF_{(1-\alpha)r+\alpha\tilde{z}} \mid \alpha \in R^1\}$ are strictly quasiconcave in α.[48]

THEOREM 4: *The following conditions on a pair of Fréchet differentiable preference functionals $V(\cdot)$ and $V^*(\cdot)$ on $D[0,M]$ with respective local utility functions $U(x;F)$ and $U^*(x;F)$ are equivalent*:

(i) *For arbitrary distributions $F(\cdot)$, $F^{**}(\cdot) \in D[0,M]$ and positive probability p, if c and c^* respectively solve $V((1-p)F^{**} + pF) = V((1-p)F^{**} + pG_c)$ and $V^*((1-p)F^{**} + pF) = V^*((1-p)F^{**} + pG_{c^*})$, then $c \leqq c^*$, (i.e. the conditional certainty equivalents for $V(\cdot)$ are never greater than the corresponding ones for $V^*(\cdot)$).*

(ii) *For all $F(\cdot) \in D[0,M]$, $U(x;F)$ is at least as concave a function of x as $U^*(x;F)$ (i.e. for all F, $U(x;F)$ is a concave transform of $U^*(x;F)$), so that if these functions are twice differentiable in x, then $-U_{11}(x;F)/U_1(x;F) \geqq -U^*_{11}(x;F)/U^*_1(x;F)$ for all x, where subscripts denote partial derivatives with respect to x.*

(iii) *If the distribution $F^*(\cdot)$ differs from $F(\cdot)$ by a simple compensated spread from the point of view of $V^*(\cdot)$ (see Section 2.1) so that $V^*(F^*) = V^*(F)$, then $V(F^*) \leqq V(F)$.*

If both individuals are diversifiers and have differentiable local utility functions, then the above conditions are equivalent to:

(iv) *For any distribution $F^{**}(\cdot) \in D[0,M]$, positive probability p, positive constant r, and nonnegative random variable $\tilde{z}$ with $E[\tilde{z}] > r$, if α and α^* yield the most preferred distributions of the form $(1-p)F^{**} + pF_{(1-\alpha)r+\alpha\tilde{z}}$ for $V(\cdot)$ and $V^*(\cdot)$ respectively, then $\alpha \leqq \alpha^*$ (i.e. the conditional demands for risky assets are never greater for $V(\cdot)$ than for $V^*(\cdot)$).*[49]

(Proof in Appendix.)

Thus the Arrow-Pratt measure of risk aversion, when applied to the local utility functions, yields a necessary and sufficient condition for one individual to

[48] This condition ensures that preferences are either (i) strictly monotonic in α or (ii) admit of a unique optimum value of α and are monotonically increasing in α below this optimum value and monotonically decreasing in α above it.

[49] Note that special cases of conditions (i) and (iv) are that the unconditional certainty equivalents are higher for $V(\cdot)$ and that the unconditional demands for the risky asset are higher for $V^*(\cdot)$, respectively.

be more risk averse than another, so that in particular, expected utility results involving the Arrow-Pratt measure as a measure of comparative risk aversion will typically apply to any pair of individuals with smooth preferences. Similarly, the Arrow-Pratt measure is evidently a sufficient tool for the analysis of comparative risk aversion in the general case. The results of this section suggest that much of the rest of expected utility analysis may be similarly generalized.[50]

4. THE SHAPE OF THE INDIVIDUAL PREFERENCE FUNCTIONAL

In this section we present a pair of hypotheses concerning the shape of the individual preference functional and show that these hypotheses are consistent with, and in many cases actually imply, each of the aspects of behavior discussed in Section 2.

4.1. *The Hypotheses*

The following hypotheses describe (I) the typical shape of a local utility function about a given initial distribution and (II) how the local utility function changes when evaluated at different initial distributions, that is, how $U(x;F)$ varies with x and F, respectively. Although they have other equivalent formulations, each is most conveniently expressed in terms of the "Arrow-Pratt" term $-U_{11}(x;F)/U_1(x;F)$ used in Theorem 4.

HYPOTHESIS I: *For any distribution* $F(\cdot)\in D[0,M]$, $-U_{11}(x;F)/U_1(x;F)$ *is a nonincreasing function of* x *over* $[0,M]$.

HYPOTHESIS II: *For any* $x\in[0,M]$ *and distributions* $F(\cdot), F^*(\cdot)\in D[0,M]$, *if* $F^*(\cdot)$ *stochastically dominates* $F(\cdot)$, *then* $-U_{11}(x;F^*)/U_1(x;F^*)\geqq -U_{11}(x;F)/U_1(x;F)$. (*That is, with respect to the partial order on* $D[0,M]$ *induced by the relation of stochastic dominance,* $-U_{11}(x;F)/U_1(x;F)$ *is "nondecreasing in* F.")

Thus, the assumptions that $R(\cdot)$ and $S(\cdot)$ are positive, increasing, exhibit declining absolute risk aversion in the Arrow-Pratt sense, and that $S(\cdot)$ is at least as concave as $R(\cdot)$, are sufficient (though not necessary) for the quadratic preference functional $\tilde{V}(\cdot)$ of equation (6) to satisfy both hypotheses.

It is important to note that Hypothesis I does not imply "decreasing absolute risk aversion in wealth" as it would if the individual were an expected utility maximizer. The willingness of an individual to insure against small risks about a certain wealth level c, for example, is given by $-U_{11}(c;G_c)/U_1(c;G_c)$, so that the effect of a wealth increase on this willingness to insure also depends on how this term is affected by changes in its second argument $G_c(\cdot)$ (see Machina **[58]**).

[50] In particular, note that the general approach and many of the specific results of this section may be extended to Fréchet differentiable preference functionals over multivariate (e.g. multi-commodity or intertemporal) distributions.

4.2. *Insurance, Lotteries, and Skewness Preference*

The types of behavior discussed in Section 2.1 all pertain to the individual's ranking of alternative shifts from an initial probability distribution over ultimate wealth. In this section we show that, when the alternative shifts are small enough, each of these types of behavior is consistent with or implied by Hypothesis I. Although Hypothesis I only suggests and is not strong enough to imply that such behavior extends to "large" shifts, the less the preference functional deviates from linearity (i.e. the less the shape of $U(\cdot;F)$ depends on F) the more this will tend to be the case as well.

As seen in Figure 4, Hypothesis I is consistent both with general risk aversion in the neighborhood of the initial distribution (Figure 4a) or with aversion to increases in risk involving low outcome values coupled with a preference for increases in risk involving high outcome values (Figure 4b).[51] Thus, since the purchase of an insurance policy against a low outcome-small probability event and the purchase of a lottery ticket yielding a small chance of a large outcome both induce small changes in the initial distribution of wealth, an individual with a local utility function as in Figure 4b would tend to purchase both, while another individual (or the same individual at another initial distribution) with local utility function as in Figure 4a would be among the class of people who purchase insurance but not lottery tickets.[52]

To determine the implications of Hypothesis I with regard to skewness preference, note that, just as in expected utility theory, Hypothesis I implies that $U_{111}(x;F)$ is positive. If $U(x;F)$ is an analytic function of x over $[0,M]$, we may

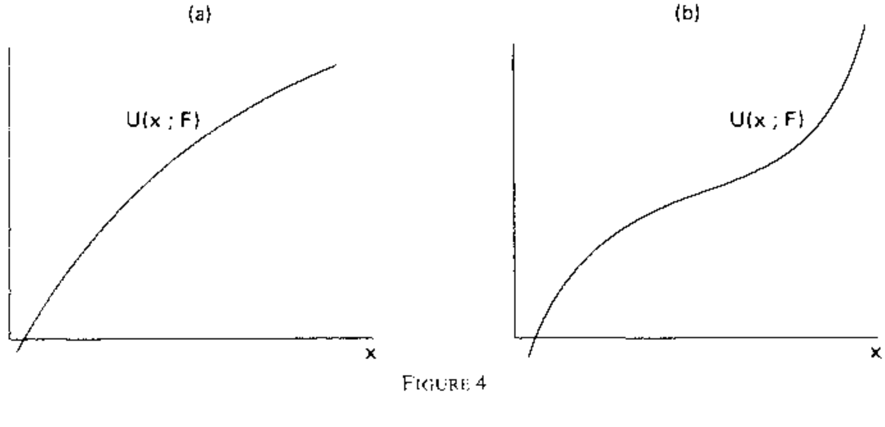

FIGURE 4

[51] There is of course a third case which satisfies Hypothesis I, namely a local utility function which is everywhere convex and has a nonincreasing Arrow-Pratt term. Since such a case implies that the individual prefers all small increases in risk, we shall not consider it further. The discussion in Section 4.4 below, however, will apply to this case as well.

[52] Note that this would be true even if the initial distribution of wealth were nondegenerate, provided that the outcome of the lottery and the event to be insured against were independent of the initial distribution of wealth.

write

$$E_{F^*}[U(x;F)] = U(\mu_{F^*};F) + \sum_{n=2}^{\infty}\left(\frac{1}{n!}\right)c_n^{F^*}\cdot U_{(n)}(\mu_{F^*};F) \equiv g(\mu_{F^*}, c_2^{F^*}, c_3^{F^*}, \ldots), \tag{10}$$

where c_n^F is the nth central moment of the distribution $F(\cdot)$ and $U_{(n)}(x;F) = d^nU(x;F)/dx^n$. When $F^*(\cdot)$ is close to $F(\cdot)$, the first order Taylor expansion of g about the moments of $F(\cdot)$ gives

$$\begin{aligned} V(F^*) - V(F) &\simeq E_{F^*}[U(x;F)] - E_F[U(x;F)] \\ &= \left[U_1(\mu_F;F) + \sum_{n=2}^{\infty}\left(\frac{1}{n!}\right)c_n^F\cdot U_{(n+1)}(\mu_F;F)\right] \\ &\quad \times(\mu_{F^*} - \mu_F) + \sum_{n=2}^{\infty}\left(\frac{1}{n!}\right)U_{(n)}(\mu_F;F)\left(c_n^{F^*} - c_n^F\right). \end{aligned} \tag{11}$$

Thus, that component of the change from $F(\cdot)$ to $F^*(\cdot)$ which is commonly taken to represent the increase in skewness, namely $c_3^{F^*} - c_3^F$, is multiplied by the positive coefficient $U_{(3)}(\mu_F;F)/3! = U_{111}(\mu_F;F)/3!$.[53]

Finally, Hypothesis I implies that the individual will have a relative preference for (equivalently, a lower aversion to) small increases in risk in the upper rather than the lower tail of an initial wealth distribution, in the sense described in Section 2.1. To see this, let $F^*(\cdot)$ differ from $F(\cdot)$ by a differential simple compensated spread. Starting from $F(\cdot)$ again and applying this same differential increase in risk to outcome values which are all uniformly greater by the (positive or negative) constant c yields the distribution $F^{**}(\cdot)$ defined by $F^{**}(x) - F(x) \equiv F^*(x-c) - F(x-c)$. Since $E_{F^*}[U(x;F)] - E_F[U(x;F)] = 0$ by assumption and $E_{F^{**}}[U(x;F)] - E_F[U(x;F)]$ equals $E_{F^*}[U(x+c;F)] - E_F[U(x+c;F)]$ by construction of $F^{**}(\cdot)$, we have that if c is positive (negative), then $F^{**}(\cdot)$ will be weakly preferred (weakly not preferred) to $F(\cdot)$, since by Hypothesis I $U(x+c;F)$ will be no less concave (at least as concave) a function of x as $U(x;F)$.

4.3. *Violations of the Independence Axiom*

In this section we offer formal characterizations of the various types of systematic violations of the independence axiom discussed in Sections 2.4 and 2.5 and demonstrate their equivalence to Hypothesis II.

The first type of behavior, mentioned in Section 2.4, was that if the two pairs of distributions (F_1, F_2) and (F_3, F_4) differ by the same "shift" (i.e. changes in the probabilities) or by scalar multiples of the same shift, so that $F_4(x) - F_3(x) \equiv \lambda\cdot$

[53] Mention should be made of Hanson and Menezes' [**41**] objections to the use of the third derivative of utility as a measure of skewness preference, which will apply to the present case as well.

$(F_2(x) - F_1(x))$ for some $\lambda > 0$, and if (F_3, F_4) respectively stochastically dominate (F_1, F_2), then the individual will rank F_3 and F_4 as if using a von Neumann-Morgenstern utility function which is no less risk averse than the one "used" to rank F_1 and F_2. Thus, if F_2 differs from F_1 by a simple compensated spread, then F_3 will be weakly preferred to F_4. On the other hand, if F_4 differs from F_3 by a simple compensated spread, then F_2 will be weakly preferred to F_1.

The second type of behavior discussed in Section 2.4 concerns the "non-independence" of the conditional certainty equivalent of a risky prospect $F(\cdot)$ in an event E with respect to the conditional distribution of wealth in $\sim E$. In particular, it was observed that stochastically dominating shifts in this latter distribution tended to lower the conditional certainty equivalent of F in E.

The third characterization of how individuals systematically violate the independence axiom, discussed in Section 2.5, was that, relative to expected utility maximization, individuals are oversensitive to changes in the probabilities of low probability-outlying events. Recall from Section 2.5 that if $x_1 < x_2 < x_3$ are three outcome levels in $[0, M]$, then any rightward shift of the probability mass of the initial distribution within the interval $[x_2, x_3]$ may be said to change the initial distribution so as to make the event x_3 "less outlying relative to the events x_1 and x_2." Similarly, any shift of mass from the interval $[x_2, x_3]$ to the interval $[x_3, M]$ may also be said to make the outcome x_3 less outlying to x_1 and x_2, since, for example, if the bulk of the initial distribution of wealth were near the level \$10,000, moving probability mass from near this level to the outcome level \$5,000,001 would make both this event *as well as the event of winning five million exactly* less outlying relative to events closer to the bulk of the distribution. Finally, a further rightward shift of mass within the interval $[x_3, M]$ may also be said to make x_3 less outlying relative to x_1 and x_2, since it changes the initial distribution in a way which makes x_3 less of a "large outcome" relative to the new initial distribution (or alternatively, it results in x_3 becoming closer to the "center" of the new distribution, and farther from the "right edge"). We thus adopt the following definition.

DEFINITION: If $x_1 < x_2 < x_3$ are three outcome levels in $[0, M]$, then any rightward (leftward) shift of the probability mass of the initial distribution within the interval $[x_2, M]$ is said to change the initial distribution so as to make x_3 *less (more) outlying relative to the outcome levels* x_1 *and* x_2. Similarly, any leftward (rightward) shift of mass within the interval $[0, x_2]$ is said to make the event x_1 *less (more) outlying relative to* x_2 *and* x_3.

Our definition of the individual's sensitivity to changes in the probabilities of events is also motivated by the discussion of Section 2.5. If $x_1 < x_2 < x_3$, define the "marginal rate of substitution between a shift of probability mass from x_2 to x_1 and a shift of probability mass from x_2 to x_3" (abbreviated MRS($x_2 \to x_1, x_2 \to x_3$)) as the amount of mass which must be shifted from x_2 to x_1 per unit amount shifted from x_2 to x_3 in order to keep the individual indifferent, when the amounts shifted are infinitesimally small. From equation (5), this marginal rate of substitution is seen to equal $(U(x_3; F) - U(x_2; F))/(U(x_2; F) - U(x_1; F))$ where F is the initial distribution.[54] Marginal rates of substitution between other pairs of shifts of mass between the three outcome levels may be defined similarly.

It is clear that for an expected utility maximizer these marginal rates of substitution will not depend on the initial distribution $F(\cdot)$. We thus say that a given change in the initial distribution makes the individual "more (less) sensitive to changes in the probability of x_3 relative to changes in the probabilities of x_1 and x_2" if the change raises (lowers) both $\text{MRS}(x_2 \to x_1, x_2 \to x_3)$ and $\text{MRS}(x_2 \to x_1, x_1 \to x_3)$, that is, if a shift of mass from either x_1 or x_2 up to x_3 now requires a larger (smaller) shift from x_2 to x_1 to leave the individual indifferent. Note that since these last two marginal rates of substitution will always differ by unity, we may define this effect in terms of its effect on $\text{MRS}(x_2 \to x_1, x_2 \to x_3)$ alone. Similarly, a change in the initial distribution makes the individual more (less) sensitive to changes in the probability of x_1 relative to changes in the probabilities of x_2 and x_3 if it raises (lowers) $\text{MRS}(x_2 \to x_3, x_2 \to x_1)$ (and hence $\text{MRS}(x_2 \to x_3, x_3 \to x_1)$), so that a shift of mass from either x_2 or x_3 down to x_1 now requires a greater (lesser) compensating shift from x_2 up to x_3. Since $\text{MRS}(x_2 \to x_3, x_2 \to x_1) = 1/\text{MRS}(x_2 \to x_1, x_2 \to x_3)$, we may combine these notions and adopt the following definition.

DEFINITION: If $x_1 < x_2 < x_3$, a given change in the initial distribution is said to both make the individual *weakly more (weakly less) sensitive to changes in the probability of x_3 relative to changes in the probabilities of x_1 and x_2* and make the individual *weakly less (weakly more) sensitive to changes in the probability of x_1 relative to changes in the probabilities of x_2 and x_3* if it preserves or raises (preserves or lowers) the value of $\text{MRS}(x_2 \to x_1, x_2 \to x_3)$.[55]

A final characterization of how the independence axiom is systematically violated in the special case of preferences over two-outcome distributions has been termed the "certainty effect" by Kahneman and Tversky [**46**] and the "common ratio" effect by MacCrimmon and Larsson [**57**] (see also Allais [3, pp. 90–92], Tversky [**98**], and Hagen [**38**][56]). This states that if an individual with initial wealth w is indifferent between a p chance of winning an (additional) amount x and a pq chance of winning y, then a pqr chance of winning y will be (weakly) preferred to a pr chance of winning x, for p, q, and $r \in [0, 1]$.[57] The

[54] Note that $\text{MRS}(x_2 \to x_1, x_2 \to x_3)$ is mathematically well defined even though x_2 may not lie in the support of $F(\cdot)$. In this case its behavioral interpretation may be seen to be the ratio of the amounts of probability mass which must be respectively shifted to x_1 and to x_3 from some outcome value which *is* in the support in order to leave the individual as well off as if the same total amount of mass had been shifted to x_2.

[55] Note that from the previous footnote this definition does not require that x_2 lie in the support of the initial distribution.

[56] Note that the "Bergen Paradox" in Hagen [**38**, pp. 278–279, 290–292] is a special case of this effect as well.

[57] Note that this condition differs from the version in Kahneman and Tversky [**46**] in that all strict inequalities have been made weak.

following theorem demonstrates that this effect is implied by each of the three types of behavior discussed in this section, which are in turn all equivalent to Hypothesis II.

THEOREM 5: *The following properties of a Fréchet differentiable preference function $V(\cdot)$ on $D[0,M]$ with twice differentiable local utility functions are equivalent*:

(i) *Hypothesis II.*

(ii) *Let $F_1(\cdot), F_2(\cdot), F_3(\cdot)$, and $F_4(\cdot) \in D[0,M]$ be such that F_3 and F_4 respectively stochastically dominate F_1 and F_2 and $F_4(x) - F_3(x) \equiv \lambda \cdot (F_2(x) - F_1(x))$ for some $\lambda > 0$. Then, if F_2 differs from F_1 by a simple compensated spread, $V(F_4) \lesseqgtr V(F_3)$. Similarly, if F_4 differs from F_3 by a simple compensated spread, then $V(F_2) \geqq V(F_1)$.*

(iii) *Given an event E with probability $p > 0$ and $F(\cdot), F^*(\cdot), F^{**}(\cdot) \in D[0,M]$ such that F^{**} stochastically dominates F^*, if c^* and c^{**} are the conditional certainty equivalents of F in E when the distribution of wealth in $\sim E$ is F^* and F^{**} respectively (i.e. if $V((1-p)F^* + pG_{c^*}) = V((1-p)F^* + pF)$ and $V((1-p)F^{**} + pG_{c^{**}}) = V((1-p)F^{**} + pF)$), then $c^* \geqq c^{**}$.*

(iv) *If x, y, and z are three outcome levels in $[0,M]$, then any change in the initial distribution of wealth which makes x more outlying relative to y and z makes the individual weakly more sensitive to changes in the probability of x relative to changes in the probabilities of y and z.*

In addition, each of these properties imply:

(v) *(The "Certainty Effect" or "Common Ratio" Effect) If, for some w, x, and $y \gtreqless 0$ and p, q, and $r \in [0,1]$ $V((1-p)G_w + pG_{w+x}) = V((1-pq)G_w + pqG_{w+y})$, then $V((1-pr)G_w + prG_{w+x}) \lesseqgtr V((1-pqr)G_w + pqrG_{w+y})$.*

(Proof in Appendix.)

Note that, unlike the examples in Section 4.2, Theorem 5 is "global" in that it applies to both small and large shifts in the distribution of wealth.

When choice is restricted to alternative distributions of the form F_{p_1,p_3}

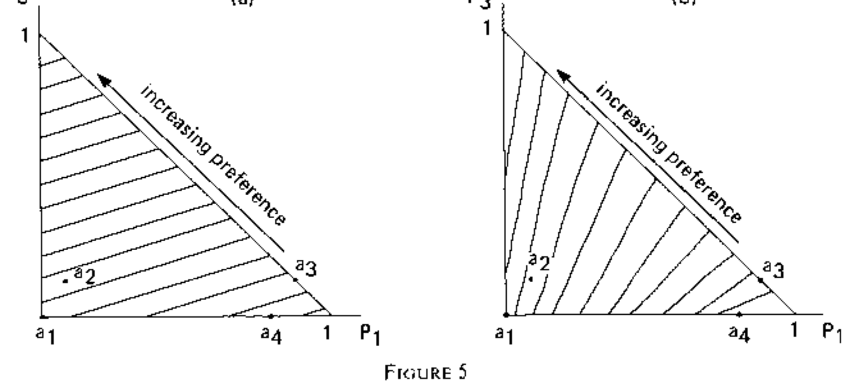

FIGURE 5

$= p_1 G_{x_1} + (1 - p_1 - p_3) G_{x_2} + p_3 G_{x_3}$ over the fixed outcomes $x_1 < x_2 < x_3$, Hypothesis II possesses a straightforward graphical interpretation in terms of indifference curves in the (p_1, p_3) plane, as in Figures 5a and 5b. It is clear that the slopes of these indifference curves, which indicate the individual's relative sensitivity to changes in p_1 versus changes in p_3, are given by $\mathrm{MRS}(x_2 \to x_3, x_2 \to x_1) = (U(x_2; F_{p_1, p_3}) - U(x_1; F_{p_1, p_3}))/(U(x_3; F_{p_1, p_3}) - U(x_2; F_{p_1, p_3}))$. Thus if the individual is an expected utility maximizer the slope will be a constant, as in Figure 5a, with a steeper slope indicating a higher level of risk aversion. However, if the individual satisfies Hypothesis II, stochastically dominating shifts in F_{p_1, p_3}, represented by upward or leftward movements in the (p_1, p_3) plane, will make the local utility function more risk averse and thus raise the slope of the indifference curves, so that the indifference curves will appear "fanned out," as in Figure 5b. The relatively steeper slopes in the region near the vertical axis than in the region near the horizontal axis illustrates the individual's greater sensitivity to changes in p_1 relative to p_3 when p_1 is small relative to p_3, and vice versa.[58]

If $x_1 = \$0$, $x_2 = \$1$ million, and $x_3 = \$5$ million, then the points corresponding to the four prospects of the Allais Paradox (Figure 3) form a parallelogram, as in Figures 5a and 5b. This illustrates why an expected utility maximizer must prefer either a_2 and a_3 if the common slope of the indifference curve is relatively flat (as in Figure 5a) or else a_1 and a_4 if it is relatively steep. Figure 5b illustrates how an individual satisfying Hypothesis II might violate the independence axiom by making the typical choices of a_1 and a_3.

4.4. *The Relative Invariance of Gambling Preferences to Initial Wealth*

In this section we demonstrate that a *fixed* preference functional $V(\cdot)$ satisfying Hypotheses I and II will tend to rank alternative gambles (expressed in terms of deviations from present wealth) relatively independently of the level of current wealth so that, unlike in the case of the expected utility model, there is no need to drop the assumption of stable preferences over $D[0, M]$ in order to accommodate the types of behavior discussed in Section 2.3.

To see this, recall that any cardinal (von Neumann-Morgenstern or local) utility function is completely characterized by its Arrow-Pratt function $-U_{11}/U_1$. Figure 6 illustrates the two alternative shapes of the local utility function of an individual with nonstochastic initial wealth level c (i.e. initial distribution $G_c(\cdot)$) together with their respective Arrow-Pratt functions, which by Hypothesis I are downward sloping. By Hypothesis II, if initial wealth increases to c^* (i.e. $G_c(\cdot)$ shifts to $G_{c^*}(\cdot)$), the Arrow-Pratt functions will shift upward, or alternatively, since they are downward sloping, shift rightward. This implies that the local utility functions will similarly shift rightward, as illustrated by the relative locations of the functions $U(\cdot; G_c)$ and $U(\cdot; G_{c^*})$ in Figure 6.

[58] This diagram clearly also applies to choices over distributions of the form $(1-p)F + p_1 G_{x_1} + (p - p_1 - p_3) G_{x_2} + p_3 G_{x_3}$ for fixed $x_1 < x_2 < x_3$, p, and $F(\cdot)$, that is, over alternative ways of distributing a probability mass of p over x_1, x_2, and x_3.

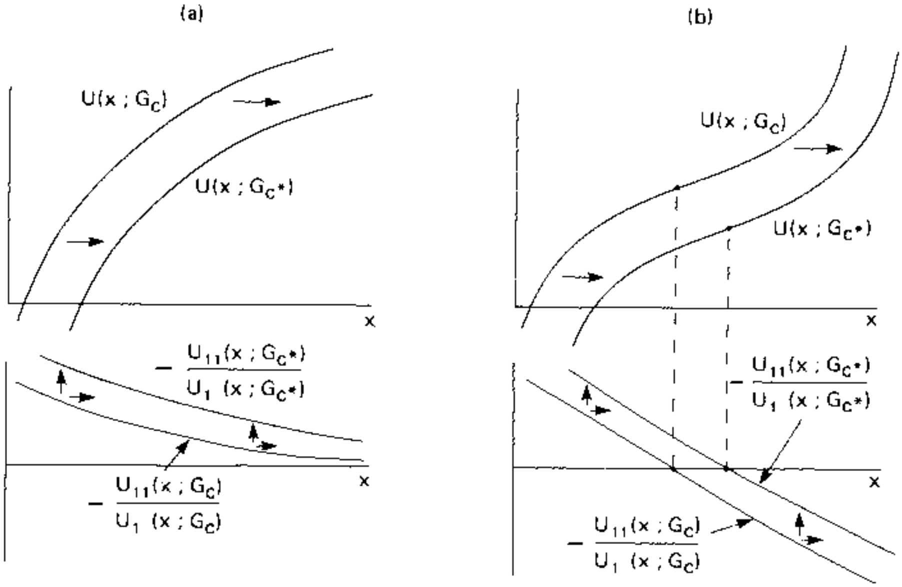

FIGURE 6

It is clear the Hypotheses I and II are not strong enough to ensure that the rightward shift in the local utility function induced by the change in initial wealth will be parallel and by the exact distance of $c^* - c$. To the extent that this happens, however, the individual's ranking of alternative gambles about initial wealth will be exactly preserved. In any event, the two hypotheses interact to ensure that preferences over gambles are less dependent upon the level of initial wealth than in the case of either of the fixed Friedman-Savage utility functions of Figures 1 and 2. In particular, it is quite possible for an individual with a fixed preference functional satisfying Hypotheses I and II to purchase insurance, purchase lottery tickets, and avoid small fair bets about all nonstochastic wealth levels. This would be the case whenever the local utility function in Figure 4b shifted so as to always keep the inflection point somewhat to the right of initial wealth (an example of such a preference functional is given in Section 4.5 below). Thus, for the purposes of explaining the behavior discussed in Section 2.3, the Markowitz assumption that shifts in initial wealth cause the entire linear preference functional to change may be replaced by the assumption that such shifts cause the *linear approximation* of a fixed nonlinear preference functional to change in the same qualitative manner. Finally, note that the two hypotheses imply that arbitrary stochastically dominating shifts in arbitrary nondegenerate initial wealth distributions will similarly cause the local utility function to shift rightward.

Besides the behavioral phenomenon of a relative invariance of gambling behavior to initial wealth, another potentially important set of behavioral observations which *cannot* be explained by Hypotheses I and II are the findings by some experimenters that individuals' expressed preferences over certain pairs of gambles can apparently be systematically reversed by increasing initial wealth by some amount and lowering each of the possible payoffs of the gambles (including zero) by the same amount, even though the two situations represent a choice over the same pair of distributions over ultimate wealth levels (see Kahneman and Tversky [**46**, p. 273]). Note that while this phenomenon is conceptually distinct from the "relative invariance . . . " phenomenon (which concerns the case when initial wealth, but not the gambles, is changed), the two are not mutually inconsistent, since an individual with an instantaneously shifting Markowitz utility function exhibits both.[59] It is, however, clearly incompatible with the existence of any fixed preference ranking over $D[0, M]$, and to the extent that it and similar observations of preference reversals and intransitivities in choice under certainty and uncertainty[60] are found to be systematic and pervasive, the behavioral model presented here must be either generalized or replaced. The point of the present section, however, is that the more extensively documented "*relative invariance . . .*" phenomenon does *not* contradict the assumption of stable preferences over $D[0, M]$, and in fact is implied by a preference ranking satisfying Hypotheses I and II.

4.5. *The St. Petersburg Paradox, Lottery Prize Structures, and the Boundedness of Utility*

In Section 2.2 it was seen that an individual with a Friedman-Savage utility function as in Figure 1 must necessarily violate each of the following "reasonable," and more to the point, commonly observed, types of behavior: (i) the amount that an individual with even minimal wealth would pay for a $1/n$ chance of winning $\$nZ$ eventually declines in n (so that lotteries will tend to have more than one prize), (ii) the individual would not forgo *any* finite sure gain to take the St. Petersburg gamble, and more generally, will assign a finite certainty equivalent to any probability distribution over nonnegative wealth levels, and (iii) there will exist a low enough positive probability p and a high enough payoff $\$C$ such that the individual will prefer a sure gain of $\$C$ to a p chance of winning any arbitrarily large prize $\$Z$.

Recall that in order to make the Friedman-Savage model compatible with these observations it was necessary to replace the terminal convex segment of the utility function with a bounded terminal concave segment.[61] In this section we

[59] It is interesting to note that although Markowitz observed that his model implied that such a change in initial wealth and the payoffs could yield an immediate preference reversal, he felt that it was "plausible to expect the chooser to act in the same manner in both situations" and sought to "resolve this dilemma" by introducing a lag between wealth changes and the shifting of the utility function [**60**, p. 155].

[60] See, for example, Kahneman and Tversky [**46**, pp. 271–273], Tversky [**97**, **98**], Grether [**34**], Grether and Plott [**35**], and the references cited in these articles.

[61] Note that this adjustment is necessary regardless of whether it is assumed that the utility function shifts when initial wealth changes.

demonstrate that, not only are these types of behavior completely consistent with Hypotheses I and II, but they are also consistent with the individual's local utility functions all possessing the terminally convex, unbounded "Friedman-Savage" shape of Figure 4b.

We demonstrate this by means of a specific example. Since each of the above types of behavior assumes that the potential outcome space is unbounded, we define the following preference functional over the space $D[0,\infty)$ of all probability distributions over the nonnegative reals:

$$\bar{V}(F) \equiv E_F[x/(1+x)] - .1 \cdot \exp(-E_F[\exp(x)]), \tag{12}$$

where $\exp(\cdot)$ is the exponential function. It is clear that for any M, the restriction of $\bar{V}(\cdot)$ to $D[0,M]$ is Fréchet differentiable with local utility function

$$\bar{U}(x;F) = x/(1+x) + .1 \cdot \exp(-E_F[\exp(z)]) \cdot \exp(x), \tag{13}$$

and with Arrow-Pratt term

$$\begin{aligned} -\bar{U}_{11}(x;F)/\bar{U}_1(x;F) &= -1 + (3+x)\Big/\Big[(1+x)\big(1+.1\cdot(1+x)^2 \\ &\qquad \times \exp(x - E_F[\exp(z)])\big)\Big]. \end{aligned} \tag{14}$$

It is not difficult to show that for any $F(\cdot)$, $-\bar{U}_{11}(x;F)/\bar{U}_1(x;F)$ is positive for low values of x, strictly decreasing in x, and eventually negative as x gets large. Thus $\bar{V}(\cdot)$ satisfies Hypothesis I over any $D[0,M]$ and (fixing F and letting M grow large enough) has local utility functions all of the shape of Figure 4b. Similarly, stochastically dominating shifts in F will raise $-\bar{U}_{11}(x;F)/\bar{U}_1(x;F)$, so that Hypothesis II is also satisfied. Since the following theorem demonstrates that $\bar{V}(\cdot)$ will prefer a certain wealth w to any other distribution in $D[0,2w]$ with the same mean, we have that an individual with this preference functional will purchase insurance, engage in lotteries, and avoid all symmetric and other small fair bets about all nonstochastic initial wealth levels. The following theorem also demonstrates that such an individual will exhibit each of the types of behavior listed at the beginning of this section.

THEOREM 6: *The preference functional* $\bar{V}(F) \equiv E_F[x/(1+x)] - .1 \cdot \exp(-E_F[\exp(x)])$ *defined over* $D[0,\infty)$ *exhibits each of the following properties*:

(i) *If* $\pi(n,w,Z)$ *is the amount that an individual with initial wealth* \$$w$ *would pay for a* $1/n$ *chance of winning an additional* \$$nZ$, *then for any* $w >$ \$.04, $\pi(n,w,Z)$ *is an eventually declining function of* n.

(ii) *For each* $F(\cdot) \in D[0,\infty)$, *there will exist a finite value* w *such that* $\bar{V}(F) < \bar{V}(G_w)$.

(iii) *For any certain initial wealth level w and probability $p < 1$, there exists a finite C such that $\bar{V}(G_{w+C}) > \bar{V}((1-p)G_w + pG_{w+Z})$ for all nonnegative Z.*[62]

(iv) *If $F^*(\cdot)$ is any nondegenerate distribution in $D[0, 2w]$ with mean w, then $\bar{V}(F^*) < \bar{V}(G_w)$, so that in particular, $\bar{V}(\cdot)$ will prefer $G_w(\cdot)$ to all other symmetric distributions in $D[0, \infty)$ with mean w.*

(*Proof in Appendix.*)

Thus, generalized expected utility analysis allows us to model a willingness to purchase lottery tickets at all wealth levels (as well as other aspects of behavior) yet avoid the adverse behavioral implications of unbounded von Neumann-Morgenstern utility functions discussed in Section 2.2. The essence of this resolution of the "boundedness of utility" debate is that the assumption of terminally convex local utility functions merely implies that the *linear approximations* to the preference functional are unbounded linear functionals, whereas the assignment of infinite certainty equivalents by an expected utility maximizer with unbounded utility, i.e. the property that $\int U(x)dF(x) = \infty = \lim_{w\to\infty} U(w)$ for some $F(\cdot)$'s in $D[0,\infty)$, follows from the fact that for such an individual the *preference functional itself* is an unbounded linear functional. Once we drop the assumption of linearity of the preference functional (i.e. the independence axiom), however, these two conditions are seen to be quite distinct, for although $\bar{V}(\cdot)$ has unbounded local utility functions, $\bar{V}(F) < 1 = \lim_{w\to\infty} \bar{V}(G_w)$ for all $F(\cdot)$ in $D[0,\infty)$.

5. CONCLUSION

5.1 *Testing Hypotheses on Preferences*

It is clear that conditions (ii), (iii), and (iv) of Theorem 5 offer three (equivalent) ways of generating further refutable implications of Hypothesis II. It is straightforward to verify, for example, that the hypothesis implies that any individual preferring a_1 to a_2 *or* a_4 to a_3 in the Allais Paradox (Figure 3) must prefer the prospect (a_5) of a .89:.11 chance of winning \$5 million or \$1 million respectively to the prospect (a_6) of a .99:.01 chance of winning \$5 million or \$0.[63] More generally, Hypothesis II would be violated by any triple of preferred prospects out of the pairs (a_1, a_2), (a_3, a_4), and (a_5, a_6) other than (a_1, a_3, a_5), (a_1, a_4, a_5), (a_2, a_3, a_5), or (a_2, a_3, a_6).

A more general approach to testing hypotheses on preferences is to parametrize $V(\cdot)$ and estimate it directly. Thus, for example, for the quadratic

[62] Note that this particular condition is slightly stronger than the corresponding condition (iii) listed at the beginning of this section.

[63] This is true since the shift from a_5 to a_6 is the same as from a_1 to a_2 and from a_4 to a_3, and since a_5 stochastically dominates a_1 and a_4.

preference functional $\tilde{V}(\cdot)$ of equation (6), we have from equation (7) that

$$(15)\qquad -\frac{\tilde{U}_{11}(x;F)}{\tilde{U}_1(x;F)} = \left[\frac{R'(x)}{R'(x)+E_F[S(z)]\cdot S'(x)}\right]\left[-\frac{R''(x)}{R'(x)}\right] + \left[\frac{E_F[S(z)]\cdot S'(x)}{R'(x)+E_F[S(z)]\cdot S'(x)}\right]\left[-\frac{S''(x)}{S'(x)}\right].$$

Equations (7) and (15) show how the properties of the preference functional depend on the properties of $R(\cdot)$ and $S(\cdot)$. Thus, if $R(\cdot)$ and $S(\cdot)$ are both positive, increasing, and concave, $\tilde{U}(\cdot;F)$ will be as well, and if, in addition, $S(\cdot)$ is more concave than $R(\cdot)$, then a stochastically dominating shift in F will, by raising $E_F[S(z)]$, raise $-\tilde{U}_{11}(x;F)/\tilde{U}_1(x;F)$. A particularly flexible four parameter functional form for $\tilde{V}(\cdot)$ can be obtained by adopting the parametrizations

$$(16)\qquad \hat{R}(x) \equiv \int_0^x \exp\left(-az-\tfrac{1}{2}bz^2\right)dz \qquad \text{and}$$

$$\hat{S}(x) \equiv \int_0^x \exp\left(-cz-\tfrac{1}{2}d\cdot z^2\right)dz,$$

which give $-\hat{R}''(x)/\hat{R}'(x) = a+bx$ and $-\hat{S}''(x)/\hat{S}'(x) = c + d\cdot x$. Thus, depending on the values of a, b, c, and d, $\hat{R}(\cdot)$ and $\hat{S}(\cdot)$ could be concave, convex, or have inflection points, and possess increasing, decreasing, or constant Arrow-Pratt terms, thus allowing for a wide range of behavior. Estimation of these parameters, say by a least squares fitting of predicted versus actual reported or observed certainty equivalents of alternative distributions, would allow for a direct test of Hypotheses I and II, as well as other hypotheses concerning the shape of $V(\cdot)$. (In this particular parametrization, Hypothesis I is valid when b and d are nonpositive, Hypothesis II is valid when $a+bx \leqq c+d\cdot x$ for all $x\in[0,M]$, and the independence axiom is equivalent to the condition that $a=c$ and $b=d$.)

Finally, while the joint consistency of Hypotheses I and II with the existence of a preference functional was demonstrated directly by the example (12), it would be useful to know whether other additional or alternative hypotheses on how $U(x;F)$ varies with x and F are similarly consistent with the existence of some $V(\cdot)$. Expressing the local utility function in "normalized form" so that $U(0;F)=0$ for all F,[64] and defining

$$\begin{aligned}(17)\qquad Q(\alpha,\beta) \equiv\ & V((1-\alpha-\beta)F+\alpha G_{z^*}+\beta G_{x^*}) \\ & - V((1-\alpha-\beta)F+\alpha G_0+\beta G_{x^*}) \\ & - V((1-\alpha-\beta)F+\alpha G_{z^*}+\beta G_0) \\ & + V((1-\alpha-\beta)F+\alpha G_0+\beta G_0),\end{aligned}$$

[64] From the discussion in footnote 43, it is clear that we may replace $U(x;F)$ by $U(x;F)-U(0;F)$ to obtain this normalized form.

we obtain

$$(18)\qquad \frac{d}{d\alpha}\left(U(x^*;(1-\alpha)F+\alpha G_{z^*})-U(x^*;(1-\alpha)F+\alpha G_0)\right)\Big|_{\alpha=0}$$

$$=\frac{d^2}{d\alpha d\beta}\left(Q(\alpha,\beta)\right)\Big|_{\alpha=\beta=0}$$

$$=\frac{d}{d\beta}\left(U(z^*;(1-\beta)F+\beta G_{x^*})-U(z^*;(1-\beta)F+\beta G_0)\right)\Big|_{\beta=0}.$$

In other words, (for $U(x;F)$ in normalized form), starting from any initial distribution $F(\cdot)$, an infinitesimal shift of probability mass from z^* to 0 will have the same effect on $U(x^*;F)$ as an equal shift of mass from x^* to 0 has on $U(z^*;F)$. While the question of sufficiency is beyond the scope of this paper, we thus have that a necessary condition for a hypothesis on how $U(x;F)$ varies with x and F to be consistent with the existence of some $V(\cdot)$ function is that it satisfy the symmetry or "integrability" condition (18).

5.2. "*Locally Utilitarian*" *Social Welfare Functionals*

Much of the analysis of Sections 3 and 4 admits of a straightforward interpretation in terms of the properties of an anonymous social welfare functional (SWF) $V(\cdot)$ defined over cumulative wealth distribution functions $F(\cdot)$ over a fixed population or measure space of agents. Because of the direct nature of the extensions, we offer neither proofs nor formal statements of theorems, but rather merely outline the types of results which may be obtained.

An important special case of such a SWF is the "utilitarian" (i.e. additively separable) functional $\int U(x)\,dF(x)$, where $U(x)$ is the social utility of an individual possessing wealth level x. It follows from Section 3.1, therefore, that if an arbitrary *non*utilitarian SWF is "smooth enough," there will exist at each wealth distribution $F(\cdot)$ a cardinal "local social utility of wealth function" $U(\cdot;F)$ such that a small change from $F(\cdot)$ to a new wealth distribution $F^*(\cdot)$ will improve social welfare (i.e. raise $V(\cdot)$) if and only if it raises aggregate local social utility, that is, if $\int U(x;F)\,dF^*(x) \geqq \int U(x;F)\,dF(x)$. It is clear from Theorem 1 that $V(\cdot)$ will satisfy the Pareto criterion (i.e. prefer an increase in any individual's wealth) if and only if all its local social utility functions are increasing in x. By Theorem 2 and the work of Atkinson [**8**], $V(\cdot)$ satisfies the Pigou-Dalton condition (i.e. is increased by costless transfers of wealth from the rich to the poor)[65] if and only if $U(x;F)$ is always concave in x. Similarly, Theorem 4 implies that the following conditions on a pair of SWF's $V(\cdot)$ and $V^*(\cdot)$ are equivalent: (i) costly (i.e. aggregate wealth lowering) transfers from richer to poorer individuals which preserve the value of $V^*(\cdot)$ will preserve or raise $V(\cdot)$;

[65]Other equivalent versions of this condition for utilitarian and nonutilitarian SWF's are discussed in Dasgupta, Sen, and Starrett [**16**] and Rothschild and Stiglitz [**75**].

(ii) the maximum acceptable cost of effecting a complete redistribution of wealth among any subgroup of society is no lower for $V(\cdot)$ than for $V^*(\cdot)$;[66] and (iii) the local social utility of wealth functions of $V(\cdot)$ are at least as concave as the corresponding ones of $V^*(\cdot)$. These equivalencies provide a natural extension of the relation "more inequality averse" to arbitrary (i.e. not necessarily utilitarian) SWF's.

Hypotheses I and II similarly admit of straightforward normative interpretations. It follows from equation (5) that the maximum acceptable proportionate deadweight loss in transferring a small amount of wealth from an individual with wealth x_b to one with wealth $x_a < x_b$ (where this loss is taken from the transferred wealth) is given by

$$(19) \qquad 1 - (U_1(x_b; F)/U_1(x_a; F)) = 1 - \exp\left(-\int_{x_a}^{x_b} (-U_{11}(x; F)/U_1(x; F))\, dx\right).$$

Hypothesis I thus implies that this maximum acceptable loss will be preserved or increased if x_a and x_b are lowered by a common amount, so that (in this sense, at least) society is at least as willing to expend resources in redistributing wealth among the poor than among the rich. Hypothesis II implies that, for fixed x_a and x_b, society's willingness to redistribute wealth between these two individuals will be preserved or increased by an improvement in the absolute wealth levels of (any or all) *other* members of society, in contrast to the utilitarian case where this willingness is independent of the wealth levels of others.[67] Together, as in Section 4.4, the two hypotheses imply that, compared with the utilitarian case of a fixed $U(\cdot)$ function, society's notions of inequality or poverty are "relative" in the sense that the local social utility function will shift rightward in response to a general increase in wealth levels.

5.3. *Related Work*

Besides the work of Kahneman and Tversky [**46**], recent years have seen a revival of interest in non-expected utility maximizing behavior. Although none of the following take the approach developed here, the reader is referred to Allais [**4**], Chew and MacCrimmon [**14**, **15**], Hagen [**38**], Kreps and Porteus [**49**], MacCrimmon and Larsson [**57**], and Selden [**84**].

University of California, San Diego

Manuscript received April, 1980; final revision received July, 1981.

[66] In both this and the previous condition, the cost of the transfer is assumed to be taken out of the transferred wealth, and not born by any other member of society.

[67] See Rothschild and Stiglitz [**75**], Sen [**85**, pp. 39–41], and the references cited there for a general discussion of the implications of additive separability in this context.

APPENDIX

LEMMA 1: *The topology of weak convergence on* $D[0, M]$ *is induced by the* L^1 *metric* $d(F^*, F) \equiv \int |F^*(x) - F(x)|\, dx$ *on* $D[0, M]$.

PROOF: Assume $F_n(x) \to F(x)$ at all continuity points x of $F(\cdot)$. Since $F(\cdot)$ is a cumulative distribution function this implies $|F_n(x) - F(x)| \to 0$ almost everywhere on $[0, M]$. Since $|F_n(x) - F(x)|$ is bounded by unity, by the Bounded Convergence Theorem (see Klambauer [**48**, pp. 59–60]) we have $d(F_n, F) = \int |F_n(x) - F(x)|\, dx \to 0$.

Conversely, let $g(\cdot)$ be an arbitrary continuous function on $[0, M]$ and ϵ an arbitrary positive number. By the Weierstrass Approximation Theorem, there exists a polynomial $\tilde{g}(\cdot)$ on $[0, M]$ such that $|\tilde{g}(x) - g(x)| < \epsilon/4$ for all $x \in [0, M]$. Also, since $\tilde{g}(\cdot)$ is a polynomial its derivative $\tilde{g}'(\cdot)$ is bounded on $[0, M]$ by some finite L. Thus, for any n, $|\int g(x)(dF_n(x) - dF(x))| < (\epsilon/2) + |\int \tilde{g}(x)(dF_n(x) - dF(x))|$, which by Lemma 2 equals $(\epsilon/2) + |\int (F_n(x) - F(x))\tilde{g}'(x)\, dx| \leqq (\epsilon/2) + L \cdot d(F_n, F)$ which, if $d(F_n, F) \to 0$, becomes less than ϵ as $n \to \infty$. *Q.E.D.*

LEMMA 2: *If* $F(\cdot)$ *is a cumulative distribution function on* $[0, M]$ *and* $g(\cdot)$ *is absolutely continuous over* $[0, M]$, *then* $\int g(x)\, dF(x) = g(M) - \int F(x)\, dg(x)$.

PROOF: Let $\{x_i\}_{i=1}^{n+1}$ be a grid on $[0, M]$ (i.e. $0 = x_1 < \cdots < x_{n+1} = M$) with norm defined as $\max_i (x_{i+1} - x_i)$. Then $g(M) = g(M) \cdot F(M) = \sum_{i=1}^{n} (F(x_{i+1})g(x_{i+1}) - F(x_i)g(x_i)) + F(0)g(0) = [F(0)g(0) + \sum_{i=1}^{n} (F(x_{i+1}) - F(x_i))g(x_i)] + [\sum_{i=1}^{n} (F(x_{i+1}) - F(x_i))(g(x_{i+1}) - g(x_i))] + [\sum_{i=1}^{n} F(x_i)(g(x_{i+1}) - g(x_i))]$. As the norm of the grid goes to zero, the first and third bracketed terms go to $\int g(x)\, dF(x)$ and $\int F(x)\, dg(x)$ respectively. By uniform continuity of $g(\cdot)$, there will exist for each positive ϵ a positive δ such that if the norm of the grid is less than δ, then $|g(x_{i+1}) - g(x_i)| < \epsilon$ for all i, so that the absolute value of the second bracketed term is less than $\epsilon \cdot \sum_{i=1}^{n} (F(x_{i+1}) - F(x_i)) \leqq \epsilon$, so that as the norm of the grid goes to zero, we obtain the desired result. *Q.E.D.*

PROOF OF THEOREM 1: Assume $F^*(\cdot)$ stochastically dominates $F(\cdot)$, and define $F(x; \alpha) \equiv \alpha F^*(x) + (1 - \alpha)F(x)$ for all $(x, \alpha) \in [0, M] \times [0, 1]$. From (9) we have

$$V(F^*) - V(F) = V(F(\cdot\,; 1)) - V(F(\cdot\,; 0))$$

$$= \int_0^1 \left[\frac{d}{d\alpha} \left(\int U(x; F(\cdot\,; \alpha^*))\, dF(x; \alpha) \right)\Big|_{\alpha^*} \right] d\alpha^*$$

$$= \int_0^1 \left[\int U(x; F(\cdot\,; \alpha^*))(dF^*(x) - dF(x)) \right] d\alpha^*.$$

However, if $U(x; F(\cdot\,; \alpha^*))$ is nondecreasing in x for all $F(\cdot\,; \alpha^*)$, then it follows from expected utility theory that the last bracketed expression will be nonnegative for all α^*, so that $V(F^*) \geqq V(F)$.

Conversely, assume that for some $\tilde{F}(\cdot) \in D[0, M]$ and $0 \leqq x^* < x^{**} \leqq M$ we have $U(x^*; \tilde{F}) > U(x^{**}; \tilde{F})$. Defining $F^*(x; \alpha) \equiv \alpha G_{x^*}(x) + (1 - \alpha)\tilde{F}(x)$ and $F^{**}(x; \alpha) \equiv \alpha G_{x^{**}}(x) + (1 - \alpha)\tilde{F}(x)$, so that $F^*(\cdot\,; 0) = F^{**}(\cdot\,; 0)$, we have from equation (8) that

$$\frac{d}{d\alpha} (V(F^*(\cdot\,; \alpha)) - V(F^{**}(\cdot\,; \alpha)))\Big|_{\alpha = 0}$$

$$= \frac{d}{d\alpha} \left[\int U(x; \tilde{F})(\alpha\, dG_{x^*}(x) + (1 - \alpha)\, d\tilde{F}(x)) - \int U(x; \tilde{F})(\alpha\, dG_{x^{**}}(x) + (1 - \alpha)\, d\tilde{F}(x)) \right]\Big|_{\alpha = 0}$$

$$= U(x^*; \tilde{F}) - U(x^{**}; \tilde{F}) > 0,$$

so that for some small positive α^* we have $V(F^*(\cdot\,; \alpha^*)) > V(F^{**}(\cdot\,; \alpha^*))$, even though $F^{**}(\cdot\,; \alpha^*)$ stochastically dominates $F^*(\cdot\,; \alpha^*)$. *Q.E.D.*

PROOFS OF THEOREMS 2 AND 3: Theorems 2 and 3 follow directly from the equivalence of conditions (i), (ii), and (iii) of Theorem 4 when $V^*(F)$ is defined to identically equal $\int x \cdot dF(x)$ (so that $U^*(x; F) \equiv x$). *Q.E.D.*

PROOF OF THEOREM 4: (i)→(ii): Assume that for some $F^{**}(\cdot) \in D[0, M]$, $U(\cdot; F^{**})$ was not at least as concave as $U^*(\cdot; F^{**})$, so that for some q and $0 \leqq x_1 < x_2 < x_3 \leqq M$,

$$0 < (U(x_2; F^{**}) - U(x_1; F^{**}))/(U(x_3; F^{**}) - U(x_1; F^{**})) < q$$
$$< (U^*(x_2; F^{**}) - U^*(x_1; F^{**}))/(U^*(x_3; F^{**}) - U^*(x_1; F^{**})) < 1.$$

Let $F(\cdot) \equiv qG_{x_3}(\cdot) + (1-q)G_{x_1}(\cdot)$. Applying equation (8) and simplifying yields

$$\frac{d}{dp}\left(V((1-p)F^{**} + pF) - V((1-p)F^{**} + pG_{x_2})\right)\Big|_{p=0}$$
$$= \int U(x; F^{**})(dF(x) - dG_{x_2}(x))$$
$$= qU(x_3; F^{**}) + (1-q)U(x_1; F^{**}) - U(x_2; F^{**}) > 0,$$

and similarly that

$$\frac{d}{dp}\left(V^*((1-p)F^{**} + pF) - V^*((1-p)F^{**} + pG_{x_2})\right)\Big|_{p=0} < 0.$$

This implies that, for some small positive p, $V((1-p)F^{**} + pF) > V((1-p)F^{**} + pG_{x_2})$ and $V^*((1-p)F^{**} + pF) < V^*((1-p)F^{**} + pG_{x_2})$, which respectively imply $c > x_2$ and $c^* < x_2$, contradicting (i).

(ii)→(iii): Let $F^*(\cdot)$ differ from $F(\cdot)$ by a simple compensated spread from the point of view of $V^*(\cdot)$, with I_L and I_R the intervals referred to in the definition of simple compensated spread (Section 2.1). Define $\phi^+(x) = \max[F^*(x) - F(x), 0]$, $\phi^-(x) = \min[F^*(x) - F(x), 0]$, and $F(x; \alpha, \beta) = F(x) + \alpha\phi^+(x) + \beta\phi^-(x)$ for all x in $[0, M]$. For $\alpha \in [0, 1]$, define $\beta(\alpha)$ as the solution to $V^*(F(\cdot; \alpha, \beta(\alpha))) = V^*(F) = V^*(F^*)$. Smoothness and strict monotonicity of $V^*(\cdot)$ ensure that $\beta(\alpha)$ is unique, increasing, differentiable, and that $\beta(0) = 0$ and $\beta(1) = 1$, so that for any α^* in $[0, 1]$,

$$\text{(A.1)} \qquad 0 = \frac{d}{d\alpha}\left(V^*(F(\cdot; \alpha, \beta(\alpha)))\right)\Big|_{\alpha^*}$$
$$= \frac{d}{d\alpha}\left(\int U^*(x; F(\cdot; \alpha^*, \beta(\alpha^*)))\, dF(x; \alpha, \beta(\alpha))\right)\Big|_{\alpha^*}$$
$$= \int U^*(x; F(\cdot; \alpha^*, \beta(\alpha^*)))[d\phi^+(x) + \beta'(\alpha^*) \cdot d\phi^-(x)].$$

Proceeding similarly with the preference functional $V(\cdot)$ yields

$$\text{(A.2)} \qquad V(F^*) - V(F) = V(F(\cdot; 1, 1)) - V(F(\cdot; 0, 0))$$
$$= V(F(\cdot; 1, \beta(1))) - V(F(\cdot; 0, \beta(0)))$$
$$= \int_0^1 \left[\frac{d}{d\alpha}\left(V(F(\cdot; \alpha, \beta(\alpha)))\right)\Big|_{\alpha^*}\right] d\alpha^*,$$

where, as in (A.1), the last bracketed term is seen to equal

$$\text{(A.3)} \qquad \int U(x; F(\cdot; \alpha^*, \beta(\alpha^*)))[d\phi^+(x) + \beta'(\alpha^*) \cdot d\phi^-(x)].$$

From (A.1) and the definitions of ϕ^+ and ϕ^- it is seen that the shift $\phi^+(\cdot) + \beta'(\alpha^*) \cdot \phi^-(\cdot) \in \Delta D[0, M]$ is a mean utility preserving increase in risk with respect to the utility function $U^*(\cdot; F(\cdot; \alpha^*, \beta(\alpha^*)))$ (see Diamond and Stiglitz [19, pp. 341–345]). Thus, by Condition (ii) of the present theorem and Theorem 3 of Diamond and Stiglitz [19], the term (A.3) is nonpositive for all α^* in [0, 1], which from (A.2) implies that $V(F^*) \leqq V(F)$.

(iii)→(i): Since $(1-p)F^{**} + pF$ differs from $(1-p)F^{**} + pG_{c^*}$ by a simple compensated spread with respect to $V^*(\cdot)$ (where $I_F = [0, c^*]$), we have $V((1-p)F^{**} + pG_{c^*}) \geqq V((1-p)F^{**} + pF) = V((1-p)F^{**} + pG_c)$, so that by monotonicity, $c^* \geqq c$.

(ii) → (iv): If for some $F^{**}(\cdot)$, p, r, and $\tilde{z}$ we have $\alpha^* < \alpha$ there would exist some $\bar{\alpha} \in (\alpha^*, \alpha)$ such that

$$\frac{d}{d\alpha}\left(V^*((1-p)F^{**} + pF_{(1-\alpha)r+\alpha\tilde{z}})\right)\Big|_{\bar{\alpha}} < 0 < \frac{d}{d\alpha}\left(V((1-p)F^{**} + pF_{(1-\alpha)r+\alpha\tilde{z}})\right)\Big|_{\bar{\alpha}}.$$

Let $F_{\tilde{z}}(\cdot)$ be the distribution of $\tilde{z}$ and

$$\bar{F}(\cdot) = (1-p)F^{**}(\cdot) + pF_{(1-\bar{\alpha})r+\bar{\alpha}\tilde{z}}(\cdot).$$

From (ii) and Pratt [**68**], we have that $x_1 < x_2$ implies $U_1(x_2; F)/U_1(x_1; F) \leqq U_1^*(x_2; F)/U_1^*(x_1; F)$ for all F in $D[0, M]$. Thus

$$\begin{aligned}
0 &< \frac{d}{d\alpha}\left(V((1-p)F^{**} + pF_{(1-\alpha)r+\alpha\tilde{z}})\right)\Big|_{\bar{\alpha}} \\
&= \frac{d}{d\alpha}\left(\int U(x; \bar{F})((1-p)\,dF^{**}(x) + p\,dF_{(1-\alpha)r+\alpha\tilde{z}}(x))\right)\Big|_{\bar{\alpha}} \\
&= p\frac{d}{d\alpha}\left(\int U((1-\alpha)r + \alpha z; \bar{F})\,dF_{\tilde{z}}(z)\right)\Big|_{\bar{\alpha}} \\
&= pU_1(r; \bar{F})\Big[\int_0^r (z-r)\big(U_1((1-\bar{\alpha})r + \bar{\alpha}z; \bar{F})/U_1(r; \bar{F})\big)\,dF_{\tilde{z}}(z) \\
&\qquad + \int_r^M (z-r)\big(U_1((1-\bar{\alpha})r + \bar{\alpha}z; \bar{F})/U_1(r; \bar{F})\big)\,dF_{\tilde{z}}(z)\Big] \\
&\leqq pU_1(r; \bar{F})\Big[\int_0^r (z-r)\big(U_1^*((1-\bar{\alpha})r + \bar{\alpha}z; \bar{F})/U_1^*(r; \bar{F})\big)\,dF_{\tilde{z}}(z) \\
&\qquad + \int_r^M (z-r)\big(U_1^*((1-\bar{\alpha})r + \bar{\alpha}z; \bar{F})/U_1^*(r; \bar{F})\big)\,dF_{\tilde{z}}(z)\Big],
\end{aligned}$$

since monotonicity of $V^*(\cdot)$ and the fact that the mean of z is greater than r imply that α^* and therefore $\bar{\alpha}$ is positive, which in turn implies that $(1-\bar{\alpha})r + \bar{\alpha}z$ will be greater than (less than) r if and only if z is greater than (less than) r. However, the positivity of the last bracketed term implies that

$$\frac{d}{d\alpha}\left(V^*((1-p)F^{**} + pF_{(1-\alpha)r+\alpha\tilde{z}})\right)\Big|_{\bar{\alpha}} > 0,$$

which is a contradiction.

(iv) → (ii): If for some $\hat{F} \in D[0, M]$ $U(\cdot; \hat{F})$ is not at least as concave as $U^*(\cdot; \hat{F})$, from Pratt [**68**] there will exist $x_1 < x_2$ and $\beta \in (0, 1)$ such that

$$U_1(x_2; \hat{F})/U_1(x_1; \hat{F}) > \beta > U_1^*(x_2; \hat{F})/U_1^*(x_1; \hat{F}),$$

so that for some small positive δ,

$$\frac{U(x_2 + \delta; \hat{F}) - U(x_2; \hat{F})}{U(x_1; \hat{F}) - U(x_1 - \beta\delta; \hat{F})} > 1 > \frac{U^*(x_2; \hat{F}) - U^*(x_2 - \delta; \hat{F})}{U^*(x_1 + \beta\delta; \hat{F}) - U^*(x_1; \hat{F})}.$$

This implies that for some positive probability p, $V(F_b) < V(F_c)$ and $V^*(F_a) > V^*(F_b)$, where

$$F_a(\cdot) = (1-p)\hat{F}(\cdot) + \frac{p}{2}G_{x_1+\beta\delta}(\cdot) + \frac{p}{2}G_{x_2-\delta}(\cdot),$$

$$F_b(\cdot) = (1-p)\hat{F}(\cdot) + \frac{p}{2}G_{x_1}(\cdot) + \frac{p}{2}G_{x_2}(\cdot), \quad \text{and}$$

$$F_c(\cdot) = (1-p)\hat{F}(\cdot) + \frac{p}{2}G_{x_1-\beta\delta}(\cdot) + \frac{p}{2}G_{x_2+\delta}(\cdot).$$

Letting

$$\alpha_a = 1 - \delta(1+\beta)/(x_2 - x_1) < \alpha_b = 1 < \alpha_c = 1 + \delta(1+\beta)/(x_2 - x_1),$$

$r = (x_1 + \beta x_2)/(1+\beta)$, and $\tilde{z}$ the random variable with distribution $F_{\tilde{z}}(\cdot) = \frac{1}{2}G_{x_1}(\cdot) + \frac{1}{2}G_{x_2}(\cdot)$, tedious algebra yields that

$$F_a = (1-p)\tilde{F} + pF_{(1-\alpha_a)r+\alpha_a\tilde{z}},$$

$$F_b = (1-p)\tilde{F} + pF_{(1-\alpha_b)r+\alpha_b\tilde{z}},$$

and

$$F_c = (1-p)\tilde{F} + pF_{(1-\alpha_c)r+\alpha_c\tilde{z}}$$

which, since both individuals are diversifiers, implies that the optimal value of α for $V(\cdot)$ is greater than α_b and that the optimal value of α for $V^*(\cdot)$ is less than α_b, contradicting (iv). *Q.E.D.*

Proof of Theorem 5: (i)$\rightarrow$(iv): Assume $x < y < z$. Then, by definition, any shift in the initial distribution which serves to make x more outlying relative to y and z must be a stochastically dominating shift, which by Hypothesis II preserves or raises the value of $-U_{11}(\xi;F)/U_1(\xi;F)$ for all ξ. From Pratt [**68**], this will preserve or lower the value of $(U(z;F) - U(y;F))/(U(y;F) - U(x;F)) = \mathrm{MRS}(y \rightarrow x, y \rightarrow z)$, which by definition makes the individual weakly more sensitive to changes in the probability of x relative to changes in the probabilities of y and z. A similar argument applies when x is greater than y and z.

(iv)$\rightarrow$(i): Assume $F^*(\cdot)$ stochastically dominates $F(\cdot)$, and let $x_1 < x_2 < x_3$ be arbitrary elements of $[0, M]$. Then the shift from F to F^* may be decomposed into a leftward shift of mass within the interval $[0, x_2]$ and a subsequent rightward shift within $[x_2, M]$ (where any mass that is ultimately shifted across x_2 is first shifted to it, then rightward from it). Since the first component of the shift makes x_1 more outlying relative to x_2 and x_3, it makes the individual weakly more sensitive to changes in the probability of x_1 relative to changes in the probabilities of x_2 and x_3, and hence preserves or lowers $\mathrm{MRS}(x_2 \rightarrow x_1, x_2 \rightarrow x_3)$. Similarly, the second component of the shift makes x_3 less outlying relative to x_1 and x_2, and hence also preserves or lowers $\mathrm{MRS}(x_2 \rightarrow x_1, x_2 \rightarrow x_3)$. Thus for the entire shift we have

$$\begin{aligned}&(U(x_3;F^*) - U(x_2;F^*))/(U(x_2;F^*) - U(x_1;F^*))\\ &\quad \leqq (U(x_3;F) - U(x_2;F))/(U(x_2;F) - U(x_1;F)).\end{aligned}$$

From Pratt [**68**], we know that if this inequality holds for arbitrary $x_1 < x_2 < x_3$, then for all x, $-U_{11}(x;F^*)/U_1(x;F^*) \geqq -U_{11}(x;F)/U_1(x;F)$.

(i)$\rightarrow$(ii): Assume F_2 differs from F_1 by a simple compensated spread. Identifying F_2 with F^*, F_1 with F, and V with V^*, define I_L, I_R, $\phi^+(\cdot)$, $\phi^-(\cdot)$, $F(\cdot;\alpha,\beta)$, and $\beta(\cdot)$ as in the proof of the implication (ii)$\rightarrow$(iii) in Theorem 4. Then, as in equation (A.1) we have

$$0 = \frac{d}{d\alpha}(V(F(\cdot;\alpha,\beta(\alpha))))\bigg|_{\alpha^*} = \int U(x;F(\cdot;\alpha^*,\beta(\alpha^*)))[d\phi^+(x) + \beta'(\alpha^*)\cdot d\phi^-(x)],$$

so that the shift $\phi^+(\cdot) + \beta'(\alpha^*)\cdot\phi^-(\cdot)$ is seen to be a mean utility preserving increase in risk with respect to the utility function $U(\cdot;F(\cdot;\alpha^*,\beta(\alpha^*)))$.

Similarly, since $F_4(\cdot) = F_3(\cdot) + \lambda\cdot(\phi^+(\cdot) + \phi^-(\cdot)) = F_3(\cdot) + \lambda\cdot\phi^+(\cdot) + \lambda\cdot\beta(1)\cdot\phi^-(\cdot)$, we have

$$\begin{aligned}V(F_4) - V(F_3) &= \int_0^1 \left[\frac{d}{d\alpha}(V(F_3 + \alpha\cdot\lambda\cdot\phi^+ + \beta(\alpha)\cdot\lambda\cdot\phi^-))\bigg|_{\alpha^*}\right] d\alpha^*\\ &= \lambda\cdot\int_0^1\left[\int U(x;F_3 + \alpha^*\lambda\phi^+ + \beta(\alpha^*)\lambda\phi^-)\right.\\ &\qquad\qquad \left.\times(d\phi^+(x) + \beta'(\alpha^*)\,d\phi^-(x))\right] d\alpha^*.\end{aligned}$$

Thus, by Hypothesis II and the argument in the proof of the implication (ii)→(iii) in Theorem 4, to prove that the above bracketed term is nonpositive for all α^* it suffices to demonstrate that the distribution $F_3+\alpha^*\lambda\phi^+ + \beta(\alpha^*)\lambda\phi^-$ stochastically dominates $F(\cdot;\alpha^*,\beta(\alpha^*))$ for all α^*.

Now, since $F_4(x)-F_2(x)=F_3(x)-F_1(x)+(\lambda-1)\phi^+(x)+(\lambda-1)\phi^-(x)$ is nonpositive for all x in $[0,M]$ as is $F_3(x)-F_1(x)$, and since α^* and $\beta(\alpha^*)$ both lie in the unit interval, we have that for all x in I_L,

$$F_3(x)+\alpha^*\lambda\phi^+(x)+\beta(\alpha^*)\lambda\phi^-(x)-F_1(x)-\alpha^*\phi^+(x)-\beta(\alpha^*)\phi^-(x)$$

$$=F_3(x)-F_1(x)+\alpha^*(\lambda-1)\phi^+(x)$$

(since $\phi^-(x)=0$ on I_L), which will be nonpositive regardless of the sign of $(\lambda-1)$. A similar argument for the case of x in I_R establishes the required stochastic dominance result.

A similar argument applies in the case when F_4 differs from F_3 by a simple compensated spread.

(ii)→(iii): Condition (iii) follows from monotonicity and Condition (ii) by defining

$$F_1=(1-p)F^*+pG_{c^*},\qquad F_2=(1-p)F^*+pF,\qquad F_3=(1-p)F^{**}+pG_{c^*},\qquad \text{and}$$

$$F_4=(1-p)F^{**}+pF.$$

(iii)→(i): The proof of this implication corresponds almost directly to the proof of the implication (i)→(ii) in Theorem 4 and is omitted.

(ii)→(v): Condition (v) is seen to be a special case of Condition (ii) when F_3, F_4, F_1, and F_2 are defined to equal the four respective arguments of $V(\cdot)$ in Condition (v), with $\lambda=1/r$ (the case when $r=0$ is trivial). *Q.E.D.*

PROOF OF THEOREM 6:(i): Define

$$F(\cdot;n,w,Z)\equiv\left(1-\frac{1}{n}\right)G_w(\cdot)+\frac{1}{n}G_{w+nZ}(\cdot).$$

Then from (8),

$$\frac{d}{dn}\left(\overline{V}(F(\cdot;n,w,Z))\right)=Z\cdot\overline{U}_1(w+nZ;F(\cdot;n,w,Z))/n$$
$$-\left(\overline{U}(w+nZ;F(\cdot;n,w,Z))-\overline{U}(w;F(\cdot;n,w,Z))\right)/n^2.$$

Substituting from (13) and rearranging gives

$$\text{(A.4)}\qquad \frac{d}{dn}\left(\overline{V}(F(\cdot;n,w,Z))\right)=\left\{-Z^2/((1+w+nZ)^2(1+w))\right\}$$
$$+\left\{.1\cdot\left(Z-\frac{1}{n}\right)\exp\left(w+nZ-E_{F(\cdot;n,w,Z)}[\exp(z)]\right)/n\right\}$$
$$+\left\{.1\cdot\exp\left(w-E_{F(\cdot;n,w,Z)}[\exp(z)]\right)/n^2\right\}.$$

Since

$$E_{F(\cdot;n,w,Z)}[\exp(z)]=\left(1-\frac{1}{n}\right)\exp(w)+\frac{1}{n}\exp(w+nZ),$$

we have that the first of the three terms on the right hand side of (A.4) goes to zero at rate $1/n^2$, and the second and third terms go to zero at a faster rate, so that for fixed w and Z, as n grows large enough, $d(\overline{V}(F(\cdot;n,w,Z)))/dn$ eventually becomes negative. It is also clear that for given w and Z there will exist a finite $n(w,Z)$ such that $n>n(w,Z)$ implies that $d(\overline{V}(F(\cdot;n,\overline{w},Z)))/dn<0$ for all $\overline{w}\in[0,w]$.

By definition, $\pi(n,w,Z)$ is the solution to $\overline{V}(F(\cdot;n,w-\pi(n,w,Z),Z))=\overline{V}(G_w)$ or, if $\overline{V}(F(\cdot;n,0,Z))\geqq\overline{V}(G_w)$, then $\pi(n,w,Z)=w$. Since $\overline{V}(F(\cdot;n,0,Z))\to 0$ as $n\to\infty$, and $\overline{V}(G_{04})>0$, provided

$w \geqq \$.04$, $\pi(n, w, Z) < w$ for large enough n. Thus for given $w^* \geqq \$.04$ and Z, we have that for large enough n^*,

$$\frac{d}{dn}(\pi(n, w^*, Z))\Big|_{n^*} = \frac{\left[(d/dn)(\bar{V}(F(\cdot\,; n, w^* - \pi(n^*, w^*, Z), Z)))|_{n^*}\right]}{\left[(d/dw)(\bar{V}(F(\cdot\,; n^*, w - \pi(n^*, w^*, Z), Z)))|_{w^*}\right]}.$$

Since the denominator in the above expression is always positive, and for $n^* \geqq n(w^*, Z)$, the numerator is always negative, we have that for fixed $w \geqq \$.04$ and Z, $d(\pi(n, w, Z))/dn$ is eventually negative for large enough n.

(ii): It is clear that for any $F(\cdot) \in D[0, \infty)$, $\bar{V}(F) < 1$, and since $\lim_{w \to \infty} \bar{V}(G_w) = 1$, for any $F(\cdot) \in D[0, \infty)$, there will exist some finite w such that $\bar{V}(F) < \bar{V}(G_w)$.

(iii): By definition,

$$\begin{aligned}\bar{V}((1-p)G_w + pG_{w+Z}) &= (w + w^2 + (p + w)Z)/((1 + w)(1 + w + Z)) \\ &\quad - .1 \cdot \exp(-(1-p)\exp(w) - p \cdot \exp(w + Z)),\end{aligned}$$

which is always strictly less than unity and approaches $(p + w)/(1 + w) < 1$ as $Z \to \infty$. Thus, since $\lim_{w \to \infty} \bar{V}(G_w) = 1$, there will exist some finite C such that $\bar{V}(G_{w+C}) > \bar{V}((1-p)G_w + pG_{w+Z})$ for all finite Z.

(iv): Defining $F(x; \alpha) \equiv \alpha F^*(x) + (1 - \alpha)G_w(x)$, we have from equation (9) that

$$\bar{V}(F^*) - \bar{V}(G_w) = \int_0^1 \left[\int_0^{2w} \bar{U}(x; F(\cdot\,; \alpha))(dF^*(x) - dG_w(x))\right] d\alpha.$$

Since the mean of F^* is w, and since from (13) or (14) it is clear that a mean preserving spread in F will increase the concavity of $\bar{U}(\cdot\,; F)$, to show that the inner integral in the above equation is always negative it suffices to show that $\bar{U}_{11}(x; G_w) < 0$ for all $x \in [0, 2w]$, or, since $\bar{U}_{111}(x; G_w)$ is easily shown to be positive, that $\bar{U}_{11}(2w; G_w) = -2/(1 + 2w)^3 + .1 \cdot \exp(2w - \exp(w)) < 0$ for all nonnegative w.

The last inequality is equivalent to $g(w) \equiv \exp(w) - 2w - 3 \cdot \ln(1 + 2w) + \ln(20) > 0$ for all $w \geqq 0$. Since $g(\cdot)$ is strictly convex and $g'(0) < 0$ and $g'(1.3) > 0$, $g(\cdot)$ will attain its minimum at $w^* \in (0, 1.3)$ where $g'(w^*) = \exp(w^*) - 2 - 6/(1 + 2w^*) = 0$. At this point, $g(w^*) = g(w^*) - g'(w^*) = -2w^* - 3 \cdot \ln(1 + 2w^*) + 6/(1 + 2w^*) + \ln(20) + 2 > 0$. *Q.E.D.*

REFERENCES

[1] ALI, M. M.: "Probability and Utility Estimates for Racetrack Bettors," *Journal of Political Economy*, 85(1977), 803–815.

[2] ALLAIS, M.: "Le Comportement de l'Homme Rationnel devant le Risque, Critique des Postulats et Axiomes de l'Ecole Americaine," *Econometrica*, 21(1953), 503–546. (Summarized version of [3].)

[3] ———: "The Foundations of a Positive Theory of Choice Involving Risk and a Criticism of the Postulates and Axioms of the American School," Part II of [5], *Expected Utility Hypotheses and the Allais Paradox*, ed. by M. Allais and O. Hagen. Dordrecht, Holland: D. Reidel Publishing Co., 1979. (Translated from *Econometrie*, Colloques Internationaux du Centre National de la Recherche Scientifique, Vol. XL, Paris, 1953, pp. 257–332.)

[4] ———: "The So-Called Allais Paradox and Rational Decisions Under Uncertainty," Part V of [5], *Expected Utility Hypotheses and the Allais Paradox*, ed. by M. Allais and O. Hagen. Dordrecht, Holland: D. Reidel Publishing Company, 1979.

[5] ALLAIS, M., AND O. HAGEN, EDS.: *Expected Utility Hypotheses and the Allais Paradox*. Dordrecht, Holland: D. Reidel Publishing Company, 1979.

[6] ARROW, K. J.: *Essays in the Theory of Risk Bearing*. Amsterdam: North Holland, 1974.

[7] ———: "The Use of Unbounded Utility Functions in Expected-Utility Maximization: Response," *Quarterly Journal of Economics*, 88(1974), 136–138.

[8] ATKINSON, A. B.: "On the Measurement of Inequality," *Journal of Economic Theory*, 2(1970), 244–263.

[9] AUMANN, R. J.: "The St. Petersburg Paradox: A Discussion of Some Recent Comments," *Journal of Economic Theory*, 14(1977), 443–445.

[10] BILLINGSLEY, P.: *Convergence of Probability Measures*. New York: John Wiley & Sons, 1968.

[11] ———: *Weak Convergence of Measures: Applications in Probability*. Philadelphia: Society for Industrial and Applied Mathematics, 1971.

[12] BINSWANGER, H. P.: "Attitudes Toward Risk: Theoretical Implications of an Experiment in Rural India," *Economic Journal*, forthcoming.

[13] BORCH, K., AND J. MOSSIN, EDS.: *Risk and Uncertainty: Proceedings of a Conference Held by the International Economic Association*. London: Macmillan and Co., 1968.

[14] CHEW, S. H., AND K. R. MACCRIMMON: "Alpha-Nu Choice Theory: A Generalization of Expected Utility Theory," University of British Columbia Faculty of Commerce and Business Administration Working Paper No. 669, July, 1979.

[15] ———: "Alpha Utility Theory, Lottery Composition and the Allais Paradox," University of British Columbia Faculty of Commerce and Business Administration Working Paper No. 686, September, 1979.

[16] DASGUPTA, P., A. SEN, AND D. STARRETT: "Notes on the Measurement of Inequality," *Journal of Economic Theory*, 6(1973), 180–187.

[17] DAVIDSON, D., P. SUPPES, AND S. SIEGEL: *Decision Making: An Experimental Approach*. Stanford, California: Stanford University Press, 1957.

[18] DEBREU, G.: *Theory of Value: An Axiomatic Analysis of General Equilibrium*. New Haven: Yale University Press, 1959.

[19] DIAMOND, P., AND J. E. STIGLITZ: "Increases in Risk and in Risk Aversion," *Journal of Economic Theory*, 8(1974), 337–360.

[20] EDEN, B.: "On Aversion to Positive Risks and Preference for Negative Risks," *Economics Letters*, 4(1979), 125–129.

[21] EDWARDS, W.: "Probability-Preferences in Gambling," *American Journal of Psychology*, 66(1953), 349–364.

[22] ———: "Probability-Preferences Among Bets With Differing Expected Values," *American Journal of Psychology*, 67(1954), 56–67.

[23] ———: "The Reliability of Probability-Preferences," *American Journal of Psychology*, 67(1954), 68–95.

[24] ———: "The Theory of Decision Making," *Psychological Bulletin*, 51(1954), 380–417.

[25] ———: "The Prediction of Decisions Among Bets," *Journal of Experimental Psychology*, 50(1955), 201–214.

[26] ———: "Behavioral Decision Theory," in *Annual Review of Psychology*, ed. by P. Farnsworth, O. McNemar, and Q. McNemar. Palo Alto: Annual Reviews, Inc., 1961.

[27] ———: "Subjective Probabilities Inferred from Decisions," *Psychological Review*, 69(1962), 109–135.

[28] FISHBURN, P.: "Binary Choice Probabilities between Gambles: Interlocking Expected Utility Models," *Journal of Mathematical Psychology*, 14(1976), 99–122.

[29] ———: "Unbounded Utility Functions in Expected Utility Theory," *Quarterly Journal of Economics*, 90(1976), 163–168.

[30] ———: "On Handa's 'New Theory of Cardinal Utility' and the Maximization of Expected Return," *Journal of Political Economy*, 86(1978), 321–324.

[31] ———: "A Probabilistic Expected Utility Theory of Risky Binary Choices," *International Economic Review*, 19(1978), 633–646.

[32] FLEMMING, J. S.: "The Utility of Wealth and the Utility of Windfalls," *Review of Economic Studies*, 36(1969), 55–66.

[33] FRIEDMAN, M., AND L. J. SAVAGE: "The Utility Analysis of Choices Involving Risk," *Journal of Political Economy*, 56(1948), 279–304. Reprinted in [91] *Readings in Price Theory*, ed. by G. Stigler and K. Boulding. Chicago: Richard D. Irwin, Inc., 1952. (See especially the correction in [91, p. 71].)

[34] GRETHER, D. M.: "Recent Psychological Studies of Behavior under Uncertainty," *American Economic Review, Papers and Proceedings*, 68(1978), 70–74.

[35] GRETHER, D. M., AND C. R. PLOTT: "Economic Theory of Choice and the Preference Reversal Phenomenon," *American Economic Review*, 69(1979), 623–638.

[36] GRIFFITH, R. M.: "Odds Adjustment by American Horse-Race Bettors," *American Journal of Psychology*, 62(1949), 290–294.

[37] HADAR, J., AND W. RUSSELL: "Rules for Ordering Uncertain Prospects," *American Economic Review*, 59(1969), 25–34.

[38] HAGEN, O.: "Towards a Positive Theory of Preferences Under Risk," in [5], *Expected Utility Hypotheses and the Allais Paradox*, ed. by M. Allais and O. Hagen. Dordrecht, Holland: D. Reidel Publishing Company, 1979, pp. 271–302.

[39] HAKANSSON, N. H.: "Friedman-Savage Utility Functions Consistent with Risk Aversion," *Quarterly Journal of Economics*, 84(1970), 472–487.

[40] HANDA, J.: "Risk, Probabilities, and a New Theory of Cardinal Utility," *Journal of Political Economy*, 85(1977), 97–122.

[41] HANSON, D. L., AND C. F. MENEZES: "On a Neglected Aspect of the Theory of Risk Aversion," *Western Economic Journal*, 9(1971), 211–217.

[42] HELPMAN, E., AND A. RAZIN: *A Theory of International Trade Under Uncertainty*. New York: Academic Press, 1978.

[43] HERSTEIN, I. N., AND J. MILNOR: "An Axiomatic Approach to Measurable Utility," *Econometrica*, 21(1953), 291–297.

[44] HIRSHLEIFER, J.: "Investment Decision Under Uncertainty: Applications of the State-Preference Approach," *Quarterly Journal of Economics*, 80(1966), 252–277.

[45] ———: *Investment, Interest, and Capital*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1970.

[46] KAHNEMAN, D., AND A. TVERSKY: "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47(1979), 263–291.

[47] KIM, Y. C.: "Choice in the Lottery-Insurance Situation: Augmented Income Approach," *Quarterly Journal of Economics*, 87(1973), 148–156.

[48] KLAMBAUER, G.: *Real Analysis*. New York: American Elsevier, 1973.

[49] KREPS, D., AND E. PORTEUS: "Temporal von Neumann-Morgenstern and Induced Preferences," *Journal of Economic Theory*, 20(1979), 81–109.

[50] KWANG, N. Y.: "Why Do People Buy Lottery Tickets? Choices Involving Risk and the Indivisibility of Expenditure," *Journal of Political Economy*, 73(1965), 530–535.

[51] LEVHARI, D., AND T. N. SRINIVASAN: "Optimal Savings Under Uncertainty," *Review of Economic Studies*, 36(1969), 153–163.

[52] LICHTENSTEIN, S.: "Bases for Preferences Among Three-Outcome Bets," *Journal of Experimental Psychology*, 69(1965), 162–169.

[53] LUCE, R., AND H. RAIFFA: *Games and Decisions: Introduction and Critical Survey*. New York: John Wiley & Sons, 1957.

[54] LUCE, R., AND P. SUPPES: "Preference, Utility, and Subjective Probability," in *Handbook of Mathematical Psychology, Volume III*, ed. by R. Luce, R. Bush, and E. Galanter. New York: John Wiley & Sons, 1965.

[55] LUENBERGER, D. G.: *Optimization by Vector Space Methods*. New York: John Wiley & Sons, 1969.

[56] MACCRIMMON, K. R.: "Descriptive and Normative Implications of the Decision-Theory Postulates," in [13], *Risk and Uncertainty: Proceedings of a Conference Held by the International Economic Association*. London: Macmillan and Co., 1968.

[57] MACCRIMMON, K. R., AND S. LARSSON: "Utility Theory: Axioms versus 'Paradoxes'," in [5], *Expected Utility Hypotheses and the Allais Paradox*, ed. by M. Allais and O. Hagen. Dordrecht, Holland: D. Reidel Publishing Company, 1979.

[58] MACHINA, M. J.: "A Stronger Characterization of Declining Risk Aversion," *Econometrica*, forthcoming.

[59] MALINVAUD, E.: "Note on von Neumann-Morgenstern's Strong Independence Axiom," *Econometrica*, 20(1952), 679.

[60] MARKOWITZ, H.: "The Utility of Wealth," *Journal of Political Economy*, 60(1952), 151–158.

[61] MARSCHAK, J.: "Rational Behavior, Uncertain Prospects, and Measurable Utility," *Econometrica*, 18(1950), 111–141. ("Errata," *Econometrica*, 18(1950), 312.)

[62] MCGLOTHLIN, W. H.: "Stability of Choices Among Uncertain Alternatives," *American Journal of Psychology*, 69(1956), 604–615.

[63] MENGER, K.: "The Role of Uncertainty in Economics," in *Essays in Mathematical Economics in Honor of Oskar Morgenstern*, ed. by M. Shubik. Princeton: Princeton University Press, 1967. (Translated from *Zeitschrift für Nationaloekonomie*, Band V, Heft 4, 1934, pp. 459–485.)

[64] MORRISON, D. G.: "On the Consistency of Preferences in Allais' Paradox," *Behavioral Science*, 12(1967), 373–383.

[65] MOSKOWITZ, H.: "Effects of Problem Representation and Feedback on Rational Behavior in Allais and Morlat-Type Problems," *Decision Sciences*, 5(1974), 225–242.

[66] MOSTELLER, F., AND P. NOGEE: "An Experimental Measurement of Utility," *Journal of Political Economy*, 59(1951), 371–404.

[67] NOGEE, P., AND B. LIEBERMAN: "The Auction Value of Certain Risky Situations," *Journal of Psychology*, 49(1960), 167–179.

[68] PRATT, J. W.: "Risk Aversion in the Small and in the Large," *Econometrica*, 32(1964), 122–136.

[69] PRESTON, M. G., AND P. BARATTA: "An Experimental Study of the Auction-Value of an Uncertain Outcome," *American Journal of Psychology*, 61(1948), 183–193.

[70] RAIFFA, H.: *Decision Analysis: Introductory Lectures on Choice Under Uncertainty*. Reading, Mass.: Addison-Wesley, 1968.

[71] ROSETT, R. N.: "The Friedman-Savage Hypothesis and Convex Acceptance Sets: A Reconciliation," *Quarterly Journal of Economics*, 81(1967), 534–535.

[72] ———: "Measuring the Perception of Risk," in [13], *Risk and Uncertainty: Proceedings of a Conference Held by the International Economic Association*. London: Macmillan and Company, 1968.

[73] ———: "Weak Experimental Verification of the Expected Utility Hypothesis," *Review of Economic Studies*, 38(1971), 481–492.

[74] ROTHSCHILD, M., AND J. STIGLITZ: "Increasing Risk: I. A Definition," *Journal of Economic Theory*, 2(1970), 225–243. ("Addendum," *Journal of Economic Theory*, 5(1972), 306.)

[75] ———: "Some Further Results on the Measurement of Inequality," *Journal of Economic Theory*, 6(1973), 188–204.

[76] ROYDEN, H.: *Real Analysis*. New York: Macmillan, 1963.

[77] RUDIN, W.: *Functional Analysis*. New York: McGraw-Hill, 1973.

[78] RUSSELL, W. R., AND T. K. SEO: "Admissible Sets of Utility Functions in Expected Utility Maximization," *Econometrica*, 46(1978), 181–184.

[79] RYAN, T. M.: "The Use of Unbounded Utility Functions in Expected Utility Maximization: Comment," *Quarterly Journal of Economics*, 88(1974), 133–135.

[80] SAMUELSON, P. A.: "Utility, Preference, and Probability," abstract of a paper given before the conference on "Les Fondements et Applications de la Théorie du Risque en Économétrie," Paris, May, 1952. Reprinted as Chapter 13 in *Collected Scientific Papers of Paul A. Samuelson*, Volume I, ed. by J. Stiglitz. Cambridge, Mass.: MIT Press, 1966.

[81] ———: "The St. Petersburg Paradox as a Divergent Double Limit," *International Economic Review*, 1(1960), 31–37.

[82] ———: "St. Petersburg Paradoxes: Defanged, Dissected, and Historically Described," *Journal of Economic Literature*, 15(1977), 24–55.

[83] SAVAGE, L. J.: *The Foundations of Statistics*. New York: Dover Publications, 1972. (Revised and enlarged version of the work originally published by John Wiley & Sons, New York, in 1954.)

[84] SELDEN, L.: "A New Representation of Preferences over 'Certain × Uncertain' Consumption Pairs: The 'Ordinal Certainty Equivalent' Hypothesis," *Econometrica*, 46(1978), 1045–1060.

[85] SEN, A.: *On Economic Inequality*. Oxford: Oxford University Press, 1973.

[86] SHAPLEY, L. S.: "The St. Petersburg Paradox: A Con Game?" *Journal of Economic Theory*, 14(1977), 439–442.

[87] ———: "Lotteries and Menus: A Comment on Unbounded Utilities," *Journal of Economic Theory*, 14(1977), 446–453.

[88] SLOVIC, P., AND A. TVERSKY: "Who Accepts Savage's Axiom?" *Behavioral Science*, 19(1974), 368–373.

[89] SPROWLS, R. C.: "Psychological-Mathematical Probability in Relationships of Lottery Gambles," *American Journal of Psychology*, 66(1953), 126–130.

[90] STIGLER, G.: "The Development of Utility Theory," Chapter 5 in *Essays in the History of Economics* by G. Stigler. Chicago: University of Chicago Press, 1965. (Reprinted from *Journal of Political Economy*, 58(1950), 307–327, 373–396.)

[91] STIGLER, G., AND K. BOULDING, EDS.: *Readings in Price Theory*. Chicago: Richard D. Irwin, Inc., 1952.

[92] STIGLITZ, J. E.: "Review of *Aspects of the Theory of Risk Bearing—Yrjö Jahnsson Lectures*, by K. J. Arrow," *Econometrica*, 37(1969), 742–743.

[93] TOBIN, J.: "Liquidity Preference as Behavior Toward Risk," *Review of Economic Studies*, 25(1957–1958), 65–86.

[94] TSAING, S. C.: "The Rationale of the Mean-Standard Deviation Analysis, Skewness Preference, and the Demand for Money," *American Economic Review*, 62(1972), 354–371.

[95] TVERSKY, A.: "Utility Theory and Additivity Analysis of Risky Choices," *Journal of Experimental Psychology*, 75(1967), 27–36.

[96] ———: "Additivity, Utility, and Subjective Probability," *Journal of Mathematical Psychology*, 4(1967), 175–201.

[97] ———: "Intransitivity of Preferences," *Psychological Review*, 76(1969), 31–48.

[98] ———: "A Critique of Expected Utility Theory: Descriptive and Normative Considerations," *Erkenntnis*, 9(1975), 163–173.

[99] VON NEUMANN, J., AND O. MORGENSTERN: *Theory of Games and Economic Behavior*. Princeton: Princeton University Press, 1944. Second Edition, 1947. Third Edition, 1953.

[100] WALLSTEN, T. S.: "Subjectively Expected Utility Theory and Subject's Probability Estimates: Use of Measurement-Free Techniques," *Journal of Experimental Psychology*, 88(1971), 31–40.

[101] YAARI, M.: "Convexity in the Theory of Choice Under Risk," *Quarterly Journal of Economics*, 79(1965), 278–290.