GENERALIZED EXPECTED UTILITY ANALYSIS AND THE NATURE OF OBSERVED VIOLATIONS OF THE INDEPENDENCE AXIOM

Mark J. Machina[1]

1. Introduction

First expressed by Allais in the early fifties, dissatisfaction with the expected utility model of individual risk taking behavior has mushroomed in recent years, as the number of papers in this volume, its predecessor (Allais & Hagen (1979)), and elsewhere[2] indicates. The nature of the current debate, i.e., whether to reject a theoretically elegant and heretofore tremendously useful descriptive model in light of accumulating evidence against its underlying assumptions, is a classic one in science, and the spur to new theoretical and empirical research which it is offering cannot help but leave economists, psychologists, and others who study this area with a better understanding of individual behavior toward risk.

In terms of its logical foundations, the expected utility model may be thought of as following from three assumptions concerning the individual's ordering of probability distributions over wealth: completeness (i.e., any two distributions can be compared), transitivity of both strict and weak preference, and the so-called "independence axiom." This latter axiom, really the cornerstone of the theory, may be stated

Mark Machina is an Associate Professor of Economics at the University of California, San Diego.

B. P. Stigum and F. Wenstøp (eds.), Foundations of Utility and Risk Theory with Applications, 263-293.

as "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." While the first two assumptions serve to imply that the individual's preferences may be represented by a real-valued maximand or "preference functional" defined over probability distributions, it is the independence axiom which gives the theory its main empirical content by placing a restriction on the functional form of the preference functional, implying that it (or some monotonic transformation of it) must be "linear in the probabilities" and hence representable as the mathematical expectation of some von Neumann-Morgenstern utility index defined over the set of pure outcomes.

Although the normative validity of the independence axiom has often been questioned in the past (see for example Allais (1952), Tversky (1975), Wold (1952), and the examples offered in Dreze (1974) and Machina (1981)), the primary form of attack on the expected utility hypothesis has been on the empirical validity of the independence axiom. Beginning with the famous example of Allais (discussed in detail below), the empirical/experimental research on the independence axiom has uncovered four types of <u>systematic</u> violations of the axiom: the "common consequence effect," the "common ratio effect" (which includes the "Bergen Paradox" and "certainty effect" as special cases), "oversensitivity to changes in small probabilities," and the "utility evaluation effect" (described below). While defenders of the expected utility model have claimed that such violations, systematic or otherwise, would disappear once the nature of such "errors" had been pointed out to subjects (e.g., Raiffa (1968, pp. 80-86), Savage (1972, pp. 102-103)), empirical tests of this assertion (MacCrimmon (1968, pp. 9-11), Slovic & Tversky (1974)) have fairly convincingly refuted it, and it is now generally acknowledged that, as a descriptive hypothesis, the independence axiom is not able to stand up to the data.

Accordingly, the defense of the expected utility model has shifted to the other two sine qua non's of a useful theory, namely analytic power and the ability to generate refutable predictions and policy implications in a wide variety of situations.[3] Expected utility supporters have pointed out that descriptive models are like lifeboats in that "you don't abandon a leaky one until something better comes along," and insist that a mere ability to rationalize "aberrant" observations is not enough for an alternative model to replace expected utility--to be acceptable, the alternative must at least approximate the analytic power and versatility of expected utility analysis. On the whole they have been correct in so arguing, as many of the alternatives which have been offered have had little predictive power, and various ones have been restricted to only pairwise choice, have implied intransitive behavior, were able to accommodate only discrete probability distributions, or even possessed the property that the individual can be led into "making book against his/herself."

The purpose of this chapter is to describe an alternative to expected utility analysis (in fact, a generalization of it) which is designed to possess the high analytic power of expected utility as well as to parsimoniously capture the nature of observed departures from the independence axiom. On the one hand, this technique, termed "generalized expected utility analysis," allows us to apply the major concepts, tools, and results of expected utility theory to the analysis of almost completely general preferences (specifically, any set of preferences which is complete, transitive, and "smooth" in the sense described below). On the other hand, however, this technique is capable of simply characterizing any additional behavioral restrictions we might feel are warranted, such as general risk aversion, declining risk aversion, comparative risk aversion between individuals, and in particular, a simple condition on preferences which serves to generate all four of the above mentioned systematic violations of the independence axiom. In addition, because of the very weak assumptions required, it turns out that many of the other alternatives and generalizations of expected utility theory which have been offered are special cases of the present analysis, which can therefore be used to derive further results in these special cases.

The following section offers a brief overview of those aspects of expected utility theory which will be relevant for the present purposes. Section 3 offers a simple graphical and algebraic description of generalized expected utility analysis, including extensions of the expected utility concepts of the "risk averse concave utility function" and the Arrow-Pratt measure of risk aversion to the general case of "smooth" preferences.[4] Section 4 offers a survey of the four known types of systematic violations of the independence axiom, as well as a description and discussion of the simple condition on preferences which serves to generate each of these four types of behavior. Section 5 offers a brief conclusion.

2. The Expected Utility Model

In this and the following sections, we adopt the standard choice-theoretic approach of assuming that the individual has a complete, transitive preference ordering over the set $D[0, M]$ of all cumulative distribution functions $F(\cdot)$ over the wealth interval $[0, M]$. As in standard consumer theory (see, for example, Debreu (1959, Ch. 4)), completeness and transitivity are sufficient to imply that we can represent the individual's ranking by some real-valued preference functional $V(\cdot)$ over $D[0, M]$, so that the probability distribution $F^*(\cdot)$ is weakly preferred to $F(\cdot)$ if and only if $V(F^*) \geq V(F)$. (In those cases when we find it useful to consider the subset $D\{x_1, \ldots, x_n\}$ of probability distributions over the payoffs $x_1 < \ldots < x_n$, we shall represent the typical distribution in $D\{x_1, \ldots, x_n\}$ by the vector of corresponding probabilities $(p_1, \ldots, p_n)$ and represent the restriction of $V(\cdot)$ to $D\{x_1, \ldots, x_n\}$ by $V(p_1, \ldots, p_n)$).

Now, if we in addition assume that the individual satisfies the independence axiom, it follows (see, e.g., Herstein & Milnor (1953)) that $V(\cdot)$ or some monotonic transformation of $V(\cdot)$ will possess the functional form $V(\cdot) \equiv \int U(x)dF(x)$ (or in the discrete case, $V(p_1, \ldots, p_n) \equiv \Sigma U(x_i, p_i)$, i.e., the mathematical expectation of the von Neumann-Morgenstern utility function $U(\cdot)$ with respect to $F(\cdot)$ (or $(p_1, \ldots, p_n)$). In other words, $V(\cdot)$ can be represented as a linear functional of $F(\cdot)$ (or in the discrete case, as a linear function of $(p_1, \ldots, p_n)$), hence the phrase that the preferences of an expected utility maximizer are "linear in the probabilities." In this case it is also clear that the distribution $F^*(\cdot)$ will be weakly preferred to $F(\cdot)$ if and only if $\int U(x)dF^*(x) \geq \int U(x)dF(x)$, or equivalently, if and only if

$$\int U(x)[dF^*(x) - dF(x)] \geq 0. \tag{1}$$

For purposes of illustration, it is useful to consider the subset $D\{x_1, x_2, x_3\}$ of all probability distributions over the wealth levels $x_1 < x_2 < x_3$ in $[0, M]$, which may be represented by the points in the unit triangle in the (p_1, p_3) plane, as in Figure 1 (with p_2 defined by $p_2 = (1 - p_1 - p_3)$. Because of the "linearity" property of expected utility maximizers, such individuals' indifference curves in this space (the solid lines in Figure 1) will be parallel straight lines, with preferred indifference curves lying to the northwest.[5] The dashed lines in the figure are what may be termed "iso-expected value loci," i.e., loci of probability distributions with the same mean. Northeast movements along such loci, since they represent changes in the distribution which preserve the mean but increase the probability of the worst and best outcomes (i.e., increase p_1 and p_3 at the expense of p_2), are seen to be precisely the set of "mean preserving spreads" in the sense of Rothschild & Stiglitz (1970). Thus, if the indifference curves are steeper than these loci, as in Figure 1, mean preserving spreads will always make the individual worse off, or in other words, the individual is risk averse. Conversely, if the indifference curves are flatter than the iso-expected value loci, the individual will be risk loving in the sense that mean preserving spreads will be preferred.

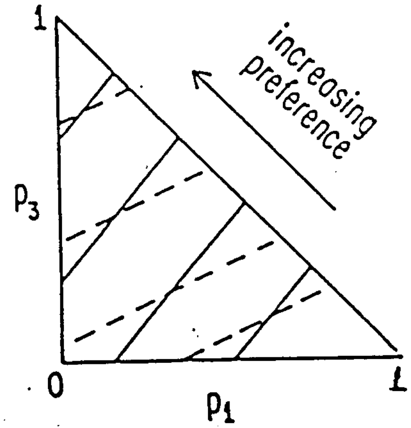

FIGURE 1

In fact, there is even a stronger sense in which the steepness of the indifference curves provides a measure of risk aversion. Solving the equation in footnote 5, we obtain that the slope of these indifference curves is equal to

$$-\frac{(U(x_3) - U(x_2)) - (U(x_2) - U(x_1))}{U(x_3) - U(x_2)} + 1. \tag{2}$$

Neglecting the addition of the constant 1, this expression (negative-the ratio of a second difference of utility to a first difference) may be thought of as the discrete analogue of the Arrow-Pratt measure $-U''(x)/U'(x)$, and indeed, Pratt (1964, Thm. 1) has shown that they are related in that the more concave the utility function, the greater the value of expression (2) for fixed x_1, x_2, and x_3. Thus, given two expected utility maximizers, the one with the steeper indifference curves will be the more risk averse over $D\{x_1, x_2, x_3\}$.

3. Generalized Expected Utility Analysis: A Brief Overview

Although there certainly have been studies which have found individual preferences over uncertain and certain prospects which violate both transitivity and completeness,[6] by far the largest and most systematic body of empirical results are those revealing systematic violations of the independence axiom. Of the three, it is in some sense fortunate that it is independence and not the other two which is most frequently violated--while dropping either transitivity or completeness would lead to a fundamental break with the traditional theory of choice, dropping independence (i.e., linearity of V(·)) amounts to simply changing the functional form of the preference functional, something which is done frequently in economic theory and econometrics.

One of the virtues of generalized expected utility analysis is that it can be developed with extremely weak assumptions on the functional form of the preference functional. Specifically, we need only assume that V(·) is a differentiable functional of F(·) (i.e., "smooth in the probabilities"), which is equivalent to assuming that indifference curves in $D\{x_1, x_2, x_3\}$ (or more generally, indifference hypersurfaces in D[0, M]) are smooth (i.e., are differentiable manifolds). Differentiability or smoothness of preferences is considered to be an extremely weak assumption in standard choice theory, and it is sufficiently weak so that many (though not all) of the functional forms which have been offered to replace expected utility are special cases of it (see below).

Algebraically, the assumption that the preference functional V(·) is differentiable in F(·) means that we can take the usual first order Taylor expansion of V(·) about any point in its domain, i.e., about any distribution $F_o(\cdot)$ in D[0, M], so that for each $F_o(\cdot)$ in D[0, M] there will exist some linear functional $\psi(\cdot;\ F_o)$ (linear in its first argument) such that

$$V(F) - V(F_o) = \psi(F - F_o;\ F_o) + o(||F - F_o||), \qquad (3)$$

where, as in standard calculus, $o(\cdot)$ denotes a function of higher order than its argument, and $||\cdot||$ is the L^1 norm, a standard measure of the "distance" between two functions.

Because $\psi(F - F_o;\ F_o)$ is linear in its first argument, it can be represented as the expectation of some function with respect to $F(\cdot) - F_o(\cdot)$, so that we may rewrite (3) as

$$V(F) - V(F_o) = \int U(x;\ F_o)[dF(x) - dF_o(x)] + o(||F - F_o||), \qquad (4)$$

where the notation $U(\cdot;\ F_o)$ is used to denote the dependence of $\Psi(\cdot;\ F_o)$, and hence its integral representation, upon the function $F_o(\cdot)$, i.e., upon the point in the domain about which we are taking the Taylor expansion. As in standard calculus, we know that for differential movements about the domain of $V(\cdot)$, (i.e., for changes from $F_o(\cdot)$ to some "very close" $F(\cdot)$), the first order or linear term in (4) will dominate the higher order term, so that the individual with preference functional $V(\cdot)$ will rank differential shifts from $F_o(\cdot)$ according to the sign of the term $\int U(x;\ F_o)[dF(x) - dF_o(x)]$. Recalling expression (1), however, we see that this is precisely the same ranking that would be used by an expected utility maximizer with a utility function $U(\cdot;\ F_o)$. Of course in some sense this is no surprise: preferences which are "smooth" (i.e., differentiable) are locally linear, and we know that in ranking probability distributions, linearity is equivalent to expected utility maximization.

Thus, even though an individual with smooth preference function $V(\cdot)$ will not necessarily satisfy the independence axiom and possesses no "global" von Neumann-Morgenstern utility function, we see that at each distribution $F_o(\cdot)$ in $D[0, M]$ there will exist a "local utility function" $U(\cdot; F_o)$ over $[0, M]$ which represents the individual's preferences at $F_o(\cdot)$. Because of the analogy between equations (1) and (4), it is clear that if $U(x; F_o)$ is increasing in x then the individual will prefer all differential first order stochastically dominating shifts from $F_o(\cdot)$,[7] and $U(x; F_o)$ will be concave in x if and only if the individual is made worse off by all differential mean preserving spreads about $F_o(\cdot)$ (i.e., is locally risk averse in the neighborhood of $F(\cdot)$).

Of course, as with any linear approximation to a differentiable function, the ranking determined by the first order linear term (i.e., by the local utility function $U(\cdot; F_o)$) will typically not correspond exactly to the ranking determined by $V(\cdot)$ over any open neighborhood of $F_o(\cdot)$ in $D[0, M]$. However, and again by analogy with standard calculus, it is possible to completely and exactly reconstruct the preference functional from knowledge of what its linear approximations (i.e., derivatives) look like at every point in the domain, by use of the Fundamental Theorem of Integral Calculus. To do this, we take any path of the form $\{F(\cdot; \alpha)|\alpha \in [0, 1]\}$ from $F_o(\cdot)$ to $F(\cdot)$ (not necessarily "near" $F_o(\cdot)$), so that $F(\cdot; 0) = F_o(\cdot)$ and $F(\cdot; 1) = F(\cdot)$, and use the fact that $V(F) - V(F_o)$ will be simply the integral of $dV(F(\cdot; \alpha))/d\alpha$ as α runs from 0 to 1. In the case of the "straight line" path $F(\cdot; \alpha) \equiv \alpha F(\cdot) + (1 - \alpha)F_o(\cdot)$, for example, we have

$$V(F) - V(F_o) = \int_0^1 \frac{dV(F(\cdot; \alpha))}{d\alpha}\, d\alpha \qquad (5)$$

$$= \int_0^1 \{\int U(x; F(\cdot; \alpha))[dF(x) - dF_o(x)]\}d\alpha,$$

since the derivative of the higher order term in (4) as α increases will be zero (see Machina (1982a) for details).

Besides yielding a way to completely reconstruct the preference functional $V(\cdot)$ from knowledge of the local utility functions, equation (5) yields insight on how generalized expected utility analysis may be used to obtain global characterizations of behavior in terms of "expected utility" type conditions on the local utility functions. For example, say that $F_1(\cdot)$ differs from $F_0(\cdot)$ by a "large" mean preserving spread. If the local utility functions $U(\cdot; F)$ are concave in x at each $F(\cdot)$, then it follows that the term in curled brackets in (5) will be nonpositive for each α, so that $V(\cdot)$ will weakly prefer $F_1(\cdot)$ to $F_0(\cdot)$. Indeed, it is shown formally in Machina (1982a) that the "expected utility" condition of concavity of (all) the local utility functions is equivalent to the individual being averse to all mean preserving spreads, or in other words, to the individual being globally risk averse.

A similar method was used in Machina (1982a) to prove two other extensions of "expected utility" analysis to the case of individuals with preference functionals which do not necessarily satisfy the independence axiom. Using straight line paths as in the previous paragraph, it is straightforward to show that the individual's preferences will exhibit "monotonicity," i.e., preference for first order stochastically dominating distributions, if and only if all the local utility functions are increasing in x. The second result extends the well known "Arrow-Pratt theorem" of comparative risk aversion: if we form the natural analogue to the Arrow-Pratt measure in our more general setting, i.e., $-U_{11}(x; F)/U_1(x; F)$ (where subscripts denote successive partial derivatives with respect to x), we have that one individual will be everywhere more risk averse than another in the standard behavioral senses (see Machina (1982a)) if and only if the "generalized Arrow-Pratt term" of the first individual is everywhere higher than that of the second, or equivalently, if and only if the first individual's local utility functions are everywhere more concave than the second's.

Note that while these types of extended expected utility theorems might seem "more complex" than those of expected utility theory since they involve checking all the local utility functions rather than a single von Neumann-Morgenstern utility function, they are in fact "less complex" in that the expected utility theorems may be thought of as derived from the more general theorems with the additional restriction that all of the local utility functions are identical.

The above algebraic arguments admit of a nice graphical interpretation in terms of the unit triangle diagram of Section 2 above. Since we are now considering preferences over the subset $D\{x_1, x_2, x_3\}$ of $D[0, M]$, we shall use the symbol $p_o = (p_{1,o}, p_{2,o}, p_{3,o})$ instead of $F_o(\cdot)$ to denote the probability distribution about which we expand the preference functional. Figure 2 illustrates the general principle that if preferences (and hence indifference curves) are smooth, then there will exist a "tangent" (i.e., linear approximating) expected utility

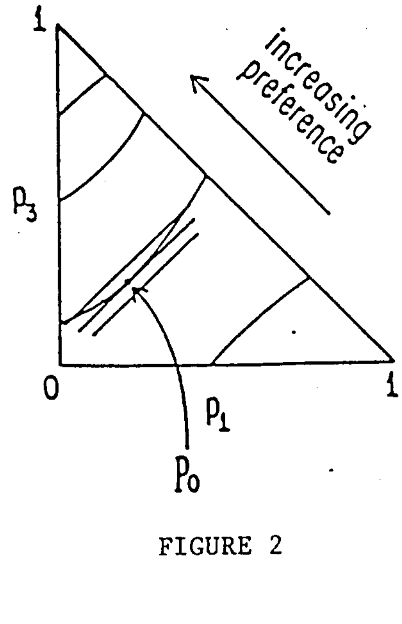

FIGURE 2

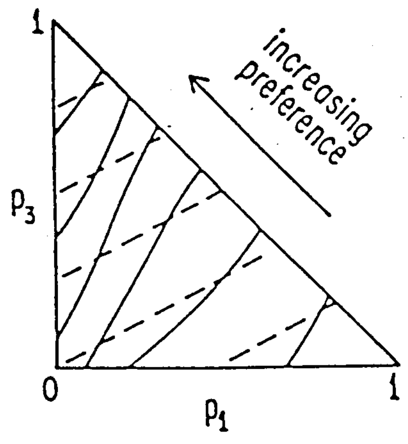

FIGURE 3

preference field to the individual's indifference curves at each distribution, as illustrated by the parallel straight lines which are tangent to the individual's actual (nonlinear) indifference curves at p_o. Figure 3 illustrates the above result that global risk aversion is equivalent to all the local utility functions being concave. Graphically, it is clear that what is necessary and sufficient for all mean preserving spreads (i.e., all northeast movements along iso-expected value lines) to make the individual worse off is not that the indifference curves necessarily be linear, but rather that they be everywhere steeper than the (dashed) iso-expected value lines. Of course, this is equivalent to the condition that the tangents to the indifference curves be everywhere steeper, which from the analysis of Section 2 is seen to be equivalent to the condition that all the local utility functions are concave in x. Finally, we could illustrate the above generalized Arrow-Pratt theorem on comparative risk aversion by a pair of nonlinear preference fields, one of whose indifference curves always intersected the other's from below (i.e., were everywhere steeper).

Having developed the above results for the case of general differentiable preference functionals, it is useful to see how they might be applied to specific special cases, i.e., to specific nonlinear functional forms. Pursuing the Taylor expansion analogy further, we see that the simplest generalization of "linearity in the probabilities" is "quadratic in the probabilities," or in other words, a functional form such as

$$V(F) \equiv \int R(x)dF(x) + \frac{1}{2}\left[\int S(x)dF(x)\right]^2, \tag{6}$$

whose local utility function can be calculated to be

$$U(x;\, F) = R(x) + S(x)\left[\int S(z)dF(z)\right]. \tag{7}$$

Thus, if R(·) and S(·) are both positive, increasing, and concave, it follows that V(·) will exhibit both monotonicity and global risk aversion, and conditions under which one preference functional of this form was everywhere more risk averse than another could similarly be determined. Table 1 presents several specific functional forms which have been suggested by researchers which are examples of smooth preference functionals, together with their calculated local utility functions.

It is clear that many more generalizations of "expected utility" type results to non-expected utility maximizers can be derived. For some examples, the reader is referred to Machina (1982a, 1982b, 1982c). We conclude this section with remarks on two issues which seem to have caused a lot of confusion in the "expected utility vs. non-expected utility" debate, namely whether non-expected utility maximizers can necessarily be tricked into "making book against themselves," and the nature of "cardinality vs. ordinality of preferences" in the context of expected utility vs. non-expected utility maximization.

There are two senses in which non-expected utility maximizers might make book against themselves (i.e., violate a preference for first order stochastic dominance in either a single choice or a sequence of choices). The first is that in certain types of non-expected utility models, most notably the "subjective expected utility" or "prospect theory" model (Edwards (1955), Kahneman & Tversky (1979)), it is necessarily true that the individual will strictly prefer some prospects to others which stochastically dominate them (see Kahneman & Tversky (1979, pp. 283-284)). Such a property of a model is clearly undesirable, and in the present author's view, makes such models unacceptable as descriptive theories (it is straightforward to show that this model is not a special case of a general differentiable preference functional). The second sense is that if an individual has a differentiable preference functional and the local utility functions are not all increasing, then the individual will prefer some distributions to others which stochastically dominate them. Of course, the analogous result is also true of expected utility maximizers: to achieve a preference for first order stochastic dominance, we must posit utility functions, von Neumann-Morgenstern or local, which are increasing in x. It is clear that the real issue is whether there can exist non-expected utility maximizing individuals who will not make book against themselves, or whether making book against oneself is an intrinsic property of non-expected utility maximizers. The answer is easy--we know from above that individuals with increasing local utility functions always prefer stochastically dominating distributions in pairwise choices, and the transitivity which follows from the maximization of $V(\cdot)$ ensures that such individuals will never violate stochastic dominance preference in a sequence of choices either.

The final issue is the apparent confusion that going from expected utility to non-expected utility involves going from "cardinal" preferences to "ordinal" preferences. This is not true. There are two related, though distinct, functions for the expected utility maximizer: the preference functional $V(\cdot)$ over $D[0, M]$ (which happens to be linear) and the von Neumann-Morgenstern utility function $U(\cdot)$ over $[0, M]$. The first of these is ordinal in that any monotonic transformation of $V(\cdot)$ will represent the same preference ranking over $D[0, M]$, and the second is cardinal in that another von Neumann-Morgenstern utility function $U^*(\cdot)$ will represent the individual's preferences if and only if $U^*(x) = aU(x) + b$ $(a > 0)$. Precisely the same is true of non-expected utility maximizers: clearly the preference functional $V(\cdot)$ of a non-expected utility maximizer is ordinal, and in Machina (1982a) it was shown that the local utility functions $U(\cdot; F)$ are cardinal in that another set of local utility functions will represent the same preferences if and only if they are a positive linear transformation of the original set. Thus, the preference functionals of all individuals, expected utility maximizing or otherwise, are always ordinal, and the utility functions, von Neumann-Morgenstern or local, are always cardinal. Whether or not the independence axiom is satisfied is irrelevant.

4. The Nature of Systematic Violations of the Independence Axiom

One of the most important points made by the defenders of expected utility theory is that dropping the independence axiom (i.e., linearity) and retaining only transitivity and completeness (and possibly smoothness) results in a model which possesses almost no predictive power. We have seen in the previous section how generalized expected utility analysis, while not requiring strong behavioral assumptions in order to apply, nevertheless still admits of refutable hypotheses such as monotonicity and risk aversion, via assumptions on the local utility functions which are analogous to the expected utility conditions. In the present section we review the evidence on the four known types of systematic violations of the independence axiom, and show that they will all follow from a single assumption on the shape of the individual preference functional V(·), which we term "Hypothesis II."[8] Thus, in addition to the usual hypotheses of monotonicity and risk aversion, generalized expected utility analysis admits of an evidently quite powerful refutable hypothesis on precisely how individuals violate the independence axiom, and one which has been substantially confirmed by the evidence so far.

4.1. The Common Consequence Effect

As an example of the first type of systematic violation of the axiom, the common consequence effect, we shall consider the first, and still most famous, specific example of this effect, namely the so-called "Allais Paradox" (see Allais (1952, p. 89), Morrison (1967), Moskowitz (1974), Raiffa (1968), and Slovic & Tversky (1974), for example). First proposed by Allais in 1952, this example consists of obtaining the subject's preference ranking over the two pairs of risky prospects

a_1:	{100% chance of \$1M	versus	a_2:	10% chance of \$5M 89% chance of \$1M 1% chance of \$ 0

and

a_3:	10% chance of \$5M 90% chance of \$ 0	versus	a_4:	11% chance of \$1M 89% chance of \$ 0

where \$1M = \$1,000,000. While it is easy to show that an expected utility maximizer would prefer either a_1 and a_4 (if $[.10U(5M) - .11U(1M) + .01U(0)] < 0$) or else a_2 and a_3 (if $[.10U(5M) - .11U(1M) + .01U(0)] > 0$), experimenters such as those listed above have repeatedly found that the modal if not majority choice of subjects has been a_1 and a_3, which violates the independence axiom.

The common consequence effect is really a generalization of the type of violation exhibited in the Allais Paradox, and involves choices between pairs of prospects of the form:

	probability	
prospect	p	1 - p
b_1	k	C*
b_2	a*	C*
b_3	k	c*
b_4	a*	c*

where a*, C*, and c* are (possibly) random prospects with C* stochastically dominating c*, and k is a sure outcome lying between the highest and lowest outcomes of a*, so that, for example, b_2 is a prospect with the same ultimate probabilities as a compound prospect yielding a p chance of a* and a 1 - p chance of C*. It is clear that an individual satisfying the independence axiom would rank b_1 and b_2 the same as b_3 and b_4: whether the "common consequence" was C* (as in the first pair) or c* (as in the second) would be "irrevelant." However, researchers such as Kahneman & Tversky (1979), MacCrimmon (1968) and MacCrimmon & Larsson (1979) as well as the five listed on the previous page have found a tendency for individuals to violate the independence axiom by preferring b_1 to b_2 and b_4 to b_3 in problems of this type (this is the same type of behavior as exhibited in the Allais Paradox, since the prospects a_1, a_2, a_3, and a_4 there correspond to b_1, b_2, b_4, and b_3, respectively, with k = C* = \$1M, c* = \$0, and a* a 10/11:1/11 chance of \$5M or \$0). In other words, the better (in the sense of stochastic dominance) the "common consequence," the more risk averse the choice (since a* is riskier than k).

4.2. The Common Ratio Effect

A second type of systematic violation of the independence axiom, the so-called "common ratio effect," also follows from an early example of Allais' (Allais (1962, p. 91)) and includes the "Bergen Paradox" of Hagen (1979) and the "certainty effect" of Kahneman & Tversky (1979) as special cases. This effect evolves rankings over pairs of prospects of the form:

$$c_1: \begin{cases} p & \text{chance of } \$X \\ 1-p & \text{chance of } \$0 \end{cases} \quad \text{versus} \quad c_2: \begin{cases} q & \text{chance of } \$Y \\ 1-q & \text{chance of } \$0 \end{cases}$$

and

$$c_3: \begin{cases} \alpha p & \text{chance of } \$X \\ 1-\alpha p & \text{chance of } \$0 \end{cases} \quad \text{versus} \quad c_4: \begin{cases} \alpha q & \text{chance of } \$Y \\ 1-\alpha q & \text{chance of } \$0 \end{cases}$$

where $p > q$, $X < Y$, and $0 < \alpha < 1$ (the term "common ratio" derives from the equality of prob(X)/prob(Y) in c_1 vs. c_2 and c_3 vs. c_4). Once again, it is clear that an individual satisfying the independence axiom would rank c_1 and c_2 the same as c_3 and c_4, however, researchers have found a systematic tendency for subjects to depart from the independence axiom by preferring c_1 to c_2 and c_4 to c_3. Thus, Kahneman & Tversky (1979) found, for example, that while 86% of their subjects preferred a .90:.10 chance of $3,000 or $0 to a .45:.55 chance of $6,000 or $0, 73% preferred a .001:.999 chance of $6,000 or $0 to a .002:.998 chance of $3,000 or $0. Besides Kahneman and Tversky, other researchers who have found this effect are Hagen (1979, pp. 285-296), MacCrimmon & Larsson (1979, pp. 350-359), and Tversky (1975).

4.3. Oversensitivity to Changes in Small Probability-outlying Events

A third type of systematic violation of the independence axiom is that, relative to the "linearity" property of expected utility, individuals tend to exhibit what may be termed an "oversensitivity to changes in the probabilities of small probability-outlying events." While the formalization of this notion requires both a precise definition of what it means for an individual to become "more sensitive" to changes in the probability of an event (relative to changes in the probabilities of certain other events) as well as what it means for an event to become "more outlying" relative to other events, we begin with an intuitive discussion of this notion, using the Allais Paradox of Section 4.1 as an example.

Note that, in the Allais example, the changes from prospects a_1 to a_2 and from a_4 to a_3 both consist of a (beneficial) shift of .10 units of probability mass from the outcome $1M to the outcome $5M and a (detrimental) shift of .01 units from $1 to $0. Since the typical individual prefers a_1 to a_2, we see that when the initial distribution is a_1, i.e., when the outcome $0 is a low probability event, the increase in its probability (at the expense of the preferred outcome $1M) is not compensated for by the beneficial shift of mass up to $5M. However, when the initial distribution is a_4, we see that the event $0 is no longer such a low probability-outlying event (since its probability is now .89) and we find that the individual is no longer as sensitive to the increase in its probability, and in the sense that the beneficial shift from \$1M to \$5M is now enough to compensate and the change to a_3 is preferred. In other words, when the initial distribution changed in a manner which made the outcome \$0 "less outlying," the individual became less sensitive to changes in its probability relative to changes in the probabilities of \$5M and \$1M.

There is an alternative way to view this example which helps bring out another aspect of the notion of "outlyingness." Note that the change in the initial distribution from a_1 to a_4 may be thought of as making the event \$5M "more outlying" relative to the events \$1M and \$0 since, although the probability of the event \$5M hasn't changed, the bulk of the distribution has moved <u>farther away</u> from the event \$5M. And in response, the individual has become more sensitive to changes in the probability of \$5M, since the beneficial increase in its probability (at the expense of \$1M) which was not enough to outweigh the detrimental shift when the initial distribution was a_1 is now enough to outweigh it when the initial distribution is a_4.

The above discussion serves as motivation for our formalitizations of the notions of "changes in sensitivity" and "outlyingness." Noting that any change in a probability distribution must consist of one or more "shifts" of probability mass from one event to another, we define the marginal rate of substitution $MRS(x_2 \rightarrow x_3, x_2 \rightarrow x_1; F)$ as the amount of probability mass which must be shifted from payoff level x_2 to x_3 per unit amount shifted from x_2 to x_1 in order to leave the individual indifferent, when the amounts shifted are infinitesimally small and the initial distribution is $F(\cdot)$. (In the following discussion, we assume $x_1 < x_2 < x_3$.) Then, the notion of increased sensitivity in the above discussion of the Allais Paradox may be formalized by saying that a change in the initial distribution $F(\cdot)$ makes the individual <u>more sensitive to changes in the probability of x_1 versus changes in the</u>

probabilities of x_2 and x_3 (and equivalently, 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 serves to raise the value of MRS($x_2 \rightarrow x_3$, $x_2 \rightarrow x_1$; F) (i.e., the individual is more sensitive to changes in the probability of x_1 if a shift of probability mass from the intermediate value x_2 to x_1 now requires more of a compensating shift of mass from x_2 up to x_3, and similarly for the case of a decreased sensitivity to changes in the probability of x_3 relative to changes in the probabilities of x_1 and x_2).

Again using the discussion of the Allais Paradox as motivation, we will say that any rightward shift of mass within the interval $[x_2, \infty)$ serves to change the initial distribution in a manner which makes the event x_3 less outlying relative to events x_1 and x_2, since rightward shifts of mass within the interval $[x_2, x_3]$ clearly move the distribution away from x_1 and x_2 and toward x_3, and rightward shifts within the interval $[x_3, \infty)$ also serve to make x_3 less of a "large" outcome relative to the bulk of the distribution, since they result in x_3 being farther from the "right edge" of the distribution. Similarly, leftward shifts of mass within the interval $(-\infty, x_2]$ serve to make the event x_1 less outlying relative to the events x_2 and x_3. Thus, our formalization of the "oversensitivity condition" is:

> any change in the initial distribution which serves to make an event more (less) outlying relative to a pair of other events serves to change the relevant marginal rate of substitution so as to make the individual more (less) sensitive to changes in the probability of that event relative to changes in the probabilities of the other two events.

While using a notion (the marginal rate of substitution) which is not typically seen in the analysis of preferences over probability distributions, the above condition is very much in the spirit of the Hicks-Allen "diminishing marginal rate of substitution" assumption of nonstochastic demand theory, in that it relates changes in a fundamental marginal rate of substitution to changes in the "current consumption bundle" (in this case, the initial distribution). Furthermore, this condition may be shown to be equivalent to the common consequence effect and to imply the common ratio effect (see Machina (1982a)), and in Section 4.5 below will be shown to possess a nice graphical interpretation in terms of the indifference curves in the unit triangle diagram.

4.4. The Utility Evaluation Effect

The final type of systematic violation of the independence axiom may be termed the "utility evaluation effect." It is well-known that there are several ways of evaluating or "assessing" the von Neumann-Morgenstern utility function of an expected utility maximizer, all of which, according to the theory, will yield the same function subject to positive linear transformations (see, for example, Farquhar (1982)). However, in actual practice different techniques have "recovered" utility functions from the same individual which differ in systematic ways.

One of the most frequently used assessment methods is termed the "fractile method" (see McCord & de Neufville (1982)). This method begins by arbitrarily defining $U(0) = 0$ and $U(M) = 1$ for some positive M, and picking some fixed probability $\bar{p}$ between zero and unity. The first step in the method then consists of determining the individual's certainty equivalent of a $\bar{p}:1 - \bar{p}$ chance of M or 0. If we term this certainty equivalent c_1, it follows from the equation $U(c_1) = \bar{p}U(M) + (1 - \bar{p})U(0)$ that $U(c_1)$ will have the value $\bar{p}$. The second and third step consists of finding the certainty equivalent c_2 of a $\bar{p}:1 - \bar{p}$ chance of c_1 and 0 (so that $U(c_2) = \bar{p}U(c_1) + (1 - \bar{p})U(0) = \bar{p}^2$) and the certainty equivalent c_3 of a $\bar{p}:1 - \bar{p}$ chance of M or c_1 (so that $U(c_3) = \bar{p}U(M) + (1 - \bar{p})U(c_1) = \bar{p} + (1 - \bar{p})\bar{p}$). Further points on the utility curve are determined by finding the certainty equivalents of a $\bar{p}:1 - \bar{p}$ chance of c_2 or 0, a $\bar{p}:1 - \bar{p}$ chance of c_1 or c_2, a $\bar{p}:1 - \bar{p}$ chance of c_3 or c_1, a $\bar{p}:1 - \bar{p}$ chance of M or c_3, etc., always interpolating by letting $\bar{p}$ be the probability of the higher of the two payoffs. Thus, if $\bar{p} = 1/2$, the first step would find that monetary value whose utility was 1/2, the second and third steps would find the values with utility levels 1/4 and 3/4, and so on through 1/8, 3/8, 5/8, 7/8, 1/16, 3/16, etc. Let $U^{\bar{p}}(\cdot)$ denote the utility function derived in this way, for a given value of $\bar{p}$.

Of course, if the individual is an expected utility maximizer, this method ought to recover the same utility function for each value of $\bar{p}$ used, i.e., the functions $U^{1/2}(\cdot)$ and $U^{1/3}(\cdot)$ ought to be identical, since both would have the same normalization $U(0) = 0$ and $U(M) = 1$. However, Karmarkar (1974) discovered an almost universal tendency for the recovered $U^{\bar{p}}(\cdot)$ to lie above the $U^{p^*}(\cdot)$ curve whenever $\bar{p}$ was higher than p^*.[9] This same effect was found (though less markedly) by McCord & de Neufville (1982) and can also be recovered from the experimental data presented by Allais (1979).[10] Once again, individuals are seen to be evidently departing from the expected utility hypothesis of linearity in a systematic manner.

4.5. Hypothesis II

The previous subsections have presented four types of systematic violations of the independence axiom that have been found by empirical researchers. Needless to say, if these four types of behavior were entirely unrelated (or even mutually contradictory), then supporters of expected utility theory would have a valid point in maintaining that any generalization of expected utility designed to accommodate them would be nothing more than an ad hoc extension of the model in each of these four directions.

However, it turns out that not only are each of the above four aspects of behavior compatible, but they all follow from a single assumption on the shape of the preference functional $V(\cdot)$. Thus, the data are telling us that not only do individuals' preferences depart from linearity, but they do so in a single systematic manner, which in addition may be modelled quite easily and which (expected utility theorists note) leads to further refutable restrictions on behavior.

As in standard calculus, one particularly compact way of specifying the nature of a nonlinearity in a preference functional is to specify how the derivative (i.e., the local utility function) of the functional varies as we move about the domain D[0, M]. Our formal hypothesis, termed "Hypothesis II,"[11] basically states that as we move from one probability distribution in D[0, M] to another which (first order) stochastically dominates it, the local utility function becomes more concave at each point x, or stated formally in terms of the Arrow-Pratt ratio $-U_{11}(x; F)/U_1(x; F)$:

<u>Hypothesis II</u>: If the distribution $F^*(\cdot)$ first order stochastically dominates $F(\cdot)$, then

$$-U_{11}(x; F^*)/U_1(x; F^*) \geq -U_{11}(x; F)/U_1(x; F)$$

for all $x \in [0, M]$.

Hypothesis II possesses a straightforward graphical interpretation in terms of the indifference curves in the unit triangle diagram. Note first that the set of all probability distributions in the triangle which stochastically dominate a given distribution corresponds to all the points which are northwest of the point representing the distribution.[12] According to Hypothesis II, therefore, the local utility functions at these northwest distributions will be more concave. However, we know from Section 3 that the more concave the (von Neumann-Morgenstern or local) utility function, the steeper the slope of the indifference curves through the point. Accordingly, Hypothesis II implies that indifference curves in the unit triangle are "fanned out" as in Figure 4, with steeper curves lying to the northwest and flatter curves lying to the southeast.

TABLE 1--LOCAL UTILITY FUNCTIONS FOR VARIOUS FUNCTIONAL FORMS OF V(·)

Mathematical Form	Reference*	V(F)	U(x; F)
Linear (i.e., expected utility)	von Neumann & Morgenstern (1944)	$\int U(x)dF(x)$	$U(x)$
Mean & variance of utility (special case of simple & general quadratic)	Allais (1952, p. 108)	$\bar{u} - \lambda\int(U(x) - \bar{u})^2 dF(x)$ $(\bar{u} = \int U(x)dF(x))$	$U(x) - \lambda U(x)^2 + 2\lambda U(x)\bar{u}$
Simple quadratic (special case of general quadratic)	Machina (1982a, p. 295)	$\int R(x)dF(x) \pm \frac{1}{2}[\int S(x)df(x)]^2$	$R(x) \pm S(x)\int S(z)dF(z)$
General quadratic	Machina (1982a, fn. 45)	$\int\int T(x, z)dF(x)dF(z)$ $(T(x, z) \equiv T(z, x))$	$2\int T(x, z)dF(z)$
First three moments of utility	Hagen (1979, p. 272)	$\bar{u} + f(s^2, m^3)$ $(\bar{u} = \int U(x)dF(x),$ $s^2 = \int(U(x) - \bar{u})^2 dF(x),$ $m^3 = \int(U(x) - \bar{u})^3 dF(x))$	$U(x) + f_1 \cdot [U(x)^2 - 2U(x)\bar{u}]$ $+ f_2 \cdot U(x)[U(x)^2 - 3U(x)\bar{u}$ $+ 3\bar{u}^2 - 3s^2]$
Rational (i.e., ratio of two linear forms)	Chew & MacCrimmon (1979) Fishburn (1981b) Bolker (1967)	$\frac{\int w(x)dF(x)}{\int \alpha(x)dF(x)}$	$\frac{w(x) - V(F)\alpha(x)^{**}}{\int \alpha(z)dF(z)}$

*The reference cited for each functional form is not necessarily the first appearance of that form, nor should it be inferred that the respective author necessarily "prefers" that form over others they may have presented. In some instances I have slightly changed the exact form as given in the reference for greater simplicity.

**I am indebted to Kenneth MacCrimmon (private correspondence) for the derivation of the local utility function of the rational form. The expression in the Table differs from his due to a difference in notation.

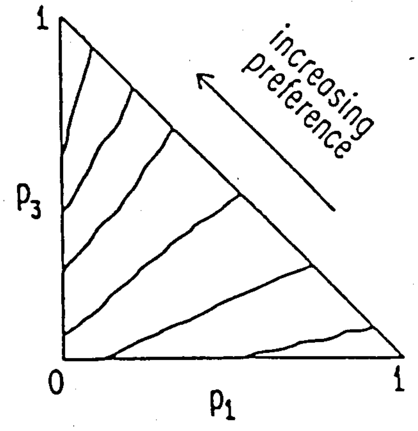

FIGURE 4

To get an idea of how Hypothesis II implies the common consequence effect, let us refer back to its general formulation (the table in Section 4.1) and consider the special case when the value k and the payoff levels of the prospects c*, C*, and a* are all elements of $\{x_1, x_2, x_3\}$ for some $x_1 < x_2 < x_3$, so that the prospects b_1, b_2, b_3, and b_4 are all in the set $D\{x_1, x_2, x_3\}$ and hence may be plotted in the unit triangle diagram. In such a case it is straightforward to show that the four prospects will always form a parallelogram with b_2 and b_4 to the northeast of b_1 and b_3 respectively, and the segment $\overline{b_1 b_2}$ parallel to and to the north and/or west of $\overline{b_3 b_4}$, e.g., as shown in Figure 5. In this case it is easy to see how the "fanning out" property of indifference curves implies by Hypothesis II would lead an individual to violate the independence axiom by preferring b_1 to b_2 and b_4 to b_3, which is precisely the common consequence effect. In Machina (1982a) it was shown that Hypothesis II is in fact <u>equivalent</u> to the common consequence effect in the more general case when c*, C*, and a* may be arbitrary (possibly continuous) prospects.

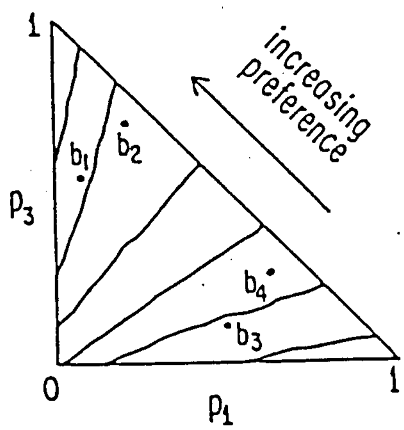

FIGURE 5

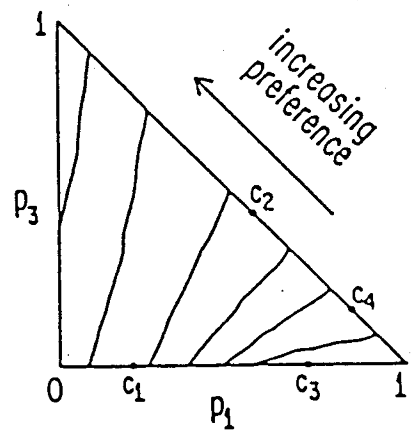

FIGURE 6

A similar graphical analysis demonstrates how Hypothesis II implies the second type of systematic violation of the independence axiom, namely the common ratio effect. Letting $x_1 = 0$, $x_2 = X$, and $x_3 = Y$ in the formulation of Section 4.2 and plotting the prospects c_1, c_2, c_3, c_4 in the unit triangle diagram, we once again find that c_2 and c_4 are northeast of c_1 and c_3 respectively and that $\overline{c_1c_2}$ is parallel to and the northeast of $\overline{c_3c_4}$, as seen in Figure 6. And similarly, it is clear how the "fanning out" property implied by Hypothesis II would lead the individual to violate the independence axiom by preferring c_1 to c_2 and c_4 to c_3, i.e., exhibit the common ratio effect.

Hypothesis II's implication that the individual will be systematically oversensitive to changes in the probabilities of low probability-outlying events may be seen quite simply from Figure 4 above. We begin by noting that, just as in nonstochastic demand theory, the marginal rate of substitution $MRS(x_2 \to x_3, x_2 \to x_1; F)$ is precisely equal to the slope of the indifference curve through the point corresponding to the distribution $F(\cdot)$ in the diagram, since rightward and upward movements in the diagram correspond to the shifts $x_2 \to x_3$ and $x_2 \to x_1$ respectively. Under the fanning out implication of Hypothesis II, we find that the individual is most sensitive to changes in the probability of x_1 relative to changes in the probabilities of x_2 and x_3 (i.e., MRS $(x_2 \to x_3, x_2 \to x_1; F)$ is the highest) near the left edge of the triangle, or in other words precisely when x_1 is a low probability event (i.e., p_1 is low). Note also that moving straight up in the triangle, which does not change p_1 but increases p_3 at the expense of p_2, also serves to make the event x_1 more outlying (since it moves probability mass further away from x_1) and indeed is seen to also increase the individual's sensitivity to changes in p_1, as measured by the slope of the indifference curves. An analogous argument applies to the individual's sensitivity to changes in p_3 relative to changes in p_1 and p_2.

Finally, we may also use the unit triangle diagram to illustrate how Hypothesis II implies the utility evaluation effect. If we were to take an individual satisfying Hypothesis II and try to "evaluate" his or her $U^{1/2}(\cdot)$ curve, the first step (as in Section 4.4) would be to determine the certainty equivalent c_1 of a 1/2:1/2 chance of M or 0. Consider now Figure 7, where we pick $x_1 = 0$, $x_2 = c_1$, and $x_3 = M$, so that the origin (i.e., the sure prospect c_1) is seen to lie on the same indifference curve as the prospect which offers a 1/2:1/2 chance of M or 0. We then find the sure amount c_2 which is indifferent to a 1/2:1/2 chance of c_1 or 0, and the amount c_3 which is indifferent to a 1/2:1/2 chance of M and c_1 (see Figure 7). These three points, with their associated $U^{1/2}(\cdot)$ values of 1/4, 1/2, and 3/4, are plotted in Figure 8 as points on the $U^{1/2}(\cdot)$ curve.

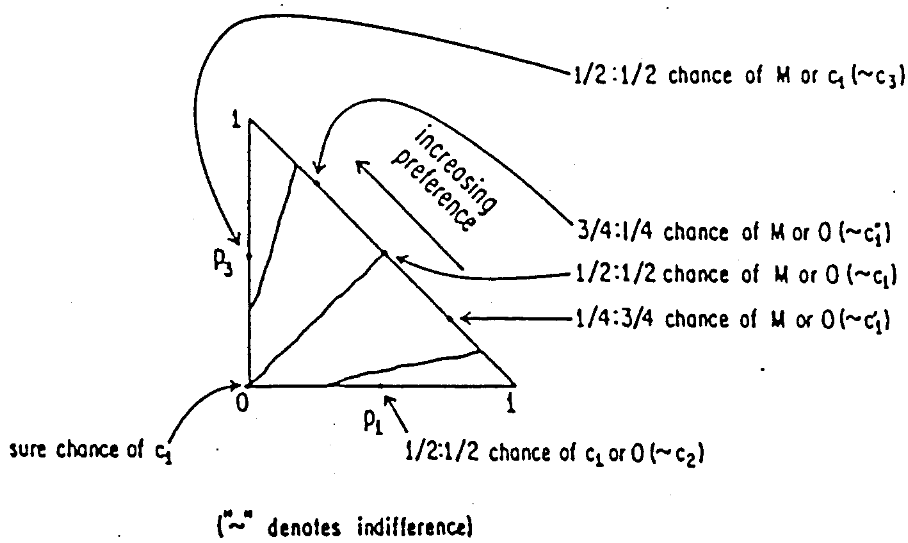

FIGURE 7

Now, to evaluate the first point on the $U^{1/4}(\cdot)$ curve, we find the certainty equivalent c_1' of a 1/4:3/4 chance of M or 0. However, if we note where this latter prospect lies in Figure 7, we see that it will be preferred to a 1/2:1/2 chance of c_1 or 0, so that its certainty equivalent c_1' will be higher than c_2. This of course implies that $U^{1/2}(\cdot)$ will attain a value of 1/4 before $U^{1/4}(\cdot)$ does, so that $U^{1/2}(\cdot)$ lies above $U^{1/4}(\cdot)$ in this region. Similarly, the first point on the $U^{3/4}(\cdot)$ curve will be the value c_1'' which is indifferent to a 3/4:1/4 chance of M or 0, and again it is seen from Figure 7 that since this prospect will be less preferred than a 1/2:1/2 chance of M or c_1, c_1'' must be less than c_3, which implies that $U^{3/4}(\cdot)$ lies above $U^{1/2}(\cdot)$ in this range (see Figure 8). This analysis may be extended to a further evaluation and comparison of the three "evaluated utility functions" in a manner which continues to exhibit the utility evaluation effect.

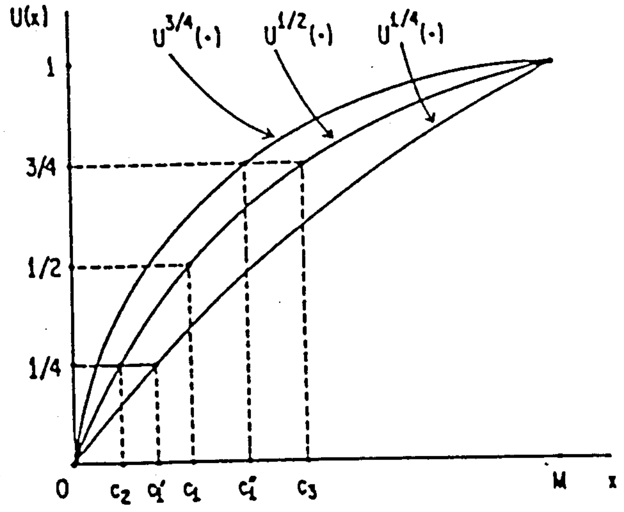

FIGURE 8

Accordingly, it is not true, as some expected utility defenders might suppose, that the violations of the independence axiom which researchers have found are random and unsystematic departures from expected utility, but rather, individuals have been found to depart from expected utility in a systematic and unified manner, as captured by Hypothesis II in general and by the "fanning out" property in the special case of preferences over three-outcome distributions.

4.6. Further Predictions and Policy Implications of Hypothesis II

It is easy to see that Hypothesis II possesses that final required property of any replacement of the expected utility hypothesis, namely the ability to generate further refutable predictions and policy implications. Of course, since each of the four types of systematic violations of expected utility discussed above is a general principle rather than a specific example, each admits of an infinite number of specific examples which serve as refutable predictions. As a new type of example, I would like to consider a problem posed by Professor Arrow in his superb and thought provoking Plenary Talk in this conference. Arrow noted that one of the canonical problems in choice under uncertainty involves the trade-off between the probability and the outcome value of an unfortunate event, and offered the specific example of an individual with initial wealth \$W facing a p probability of a loss of \$X (with a 1 - p probability of no loss). A natural question to ask here is how does the individual's marginal rate of substitution between p and X depend upon their existing values? Defining expected utility $\phi(p, X; W) \equiv pU(W - X) + (1 - p)U(W)$, we get that this marginal rate of substitution is

$$MRS_{p,\,X} = \frac{dp}{dX}\bigg|_{\phi} = \frac{-pU'(W - X)}{U(W) - U(W - X)} \tag{8}$$

In his talk, Arrow noted that this expected utility formulation implied a possibly quite useful restriction on behavior, namely that, fixing X and W, the marginal rate of substitution between p and X is proportional to p, i.e., to the probability of the unfortunate event. He quite rightly noted that it would be possible to exploit this property to make important predictions as well as policy suggestions, say in determining the trade-off between the probability and severeness of a nuclear accident, and also noted that any acceptable alternative to expected utility would have to possess this same type of ability.

To see how generalized expected utility analysis, and more particularly Hypothesis II, might be applied to this problem, we replace the expected utility maximand $\phi(p, X; W)$ with the more general maximand $V(pG_{W-X} + (1 - p)G_W)$, where G_c stands for the distribution with unit mass at c, so that $pG_{W-X} + (1 - p)G_W$ represents the distribution in question. We then have from equation (4) that

$$\text{MRS}_{p,\,X} = \frac{dp}{dX}\Big|_V \tag{9}$$

$$= \frac{-pU_1(W - X;\; pG_{W-X} + (1-p)G_W)}{U(W;\; pG_{W-X} + (1-p)G_W) - U(W - X;\; pG_{W-X} + (1-p)G_W)} =$$

(after some manipulation)

$$= -p\Big[\int_{W-X}^{W} \exp\Big[-\int_{W-X}^{z} \Big\{-\frac{U_{11}(\omega;\; pG_{W-X} + (1-p)G_W)}{U_1(\omega;\; pG_{W-X} + (1-p)G_W)}\Big\}\, d\omega\Big]dz\Big]^{-1}.$$

As usual in generalized expected utility analysis, we see the formal analogy with the expected utility case: the marginal rate of substitution in (9) is identical to that in (8) with the von Neumann-Morgenstern utility function $U(\cdot)$ replaced by the local utility function $U(\cdot;\, F)$ when $F = pG_{W-X} + (1 - p)G_W$. However, since the local utility function in (9) now depends on the precise distribution $pG_{W-X} + (1 - p)G_W$, the marginal rate of substitution is no longer strictly proportional to p as before. However, this is not to say that Hypothesis II is without implications in this case. Noting that an increase in p induces a first order stochastic worsening of the distribution $pG_{W-X} + (1 - p)G_W$, we see that under Hypothesis II an increase in p will lower the term in curled brackets in (9) (the Arrow-Pratt term) for each value of ω, so that Hypothesis II implies that the marginal rate of substitution between p and X varies less than proportionately with p. The replacement of the expected utility prediction of exact proportionality with a weak inequality on proportionality reflects the fact that Hypothesis II is a weak inequality which includes the expected utility case (i.e., the independence axiom) as a borderline case, just as, geometrically, "fanning out" includes parallel linear indifference curves as a borderline case. Nevertheless, weak inequalities are still refutable restrictions on behavior (we use them all the time in economics) and this result is clearly not without policy implications which, if not as strong as the ones generated by expected utility, are at least more accurately tied to what we have observed about individuals' actual preferences. While this is just a single example, it should be clear that Hypothesis II can be used to derive other important behavioral predictions and policy implications. .

5. Conclusion

Defenders of the expected utility approach are quite correct in insisting that any alternative to expected utility not only be consistent with the data, but also be at least on the order of elegance of the expected utility theory, and capable of easily derived behavioral restrictions and implications for policy analysis. The technique of generalized expected utility analysis seems to fit these requirements. Specifically,

> while making virtually no requisite assumptions on preferences other than completeness, transitivity, and smoothness, it allows us to retain the elegant set of concepts, tools, and techniques of expected utility analysis,
>
> it admits of refutable restrictions on preferences and hence on behavior, with the concepts of monotonicity and risk aversion, for example, modeled almost exactly as in expected utility analysis, and
>
> it admits of a restriction (Hypothesis II) which implies the four known types of observed systematic violations of the independence axiom, and which generates both additional refutable behavioral predictions as well as policy implications.

Whether the future will yield empirical observations which contradict Hypothesis II, or even the underlying assumption of smooth preferences, is really not the issue at hand.[13] The present point is that generalized expected utility analysis seems to offer a theoretically powerful and empirically supported generalization of the expected utility model. Indeed, if generalized expected utility analysis and other related models lead to the type of empirical work which will require still newer models to replace them, they will have served us well.

NOTES

1. I am indebted to Maurice Allais, Kenneth Arrow, John Harsanyi, and Ed McClennen for discussions of this material during the conference, and to Beth Hayes, Joel Sobel, and Halbert White for helpful comments on the manuscript. All errors and opinions, however, are my own.
2. See for example Chew & MacCrimmon (1979), Fishburn (1981a, 1981b), Handa (1977), and Kahneman & Tversky (1979).
3. Of course, any comparison of the refutable implications of two competing models should be followed immediately by a discussion of which of these implications have and have not in fact been refuted.
4. For a more complete and rigorous treatment of much of the material in Sections 3 and 4, see Machina (1982a, 1982b, 1982c).
5. The indifference curves here are the loci of solutions to the equation $p_1U(x_1) + (1 - p_1 - p_3)U(x_2) + p_3U(x_3) = k$ for different values of the constant k. Northwest movements make the individual better off since they consist of either increases in p_3 at the expense of p_2, increases in p_2 at the expense of p_1, or a combination of the two.
6. See for example Kahneman & Tversky (1979, pp. 271-273), Tversky (1967, 1975), Grether (1978), and Grether & Plott (1979).
7. See Hadar & Russell (1969) for the definition of first order stochastic dominance.
8. "Hypothesis I" is a separate hypothesis on the typical shape of the local utility function which, in conjunction with Hypothesis II, serves to generate behavior of the type observed by Friedman & Savage (1948) and Markowitz (1952) (see Machina (1982a)).
9. Of Karmarker's four subjects, three exhibited fitted $U^p(\cdot)$ curves which strictly and markedly increased with p. The fourth ("Subject B") exhibited $U^{9/10}(\cdot)$ and $U^{3/4}(\cdot)$ curves which were both

above the $U^{1/2}(\cdot)$ curve, but which crossed each other at one point. Since the curves of this subject were much closer to each other than the curves of the other subjects, it is possible that this crossing is due to the slightly random character of responses which is typically found in studies of this type.

10. McCord & de Neufville found that the greater majority of their subjects exhibited $U^{1/4}(\cdot)$ curves which were below their $U^{1/2}(\cdot)$ in the region where the curves had a value of 1/4. However, an equal number of their subjects had $U^{3/4}(\cdot)$ curves above and below their $U^{1/2}(\cdot)$ curves, indicating no average departure from linearity in either direction in this region. McCord & de Neufville also found that whether the $U^{1/4}(\cdot)$ and $U^{3/4}(\cdot)$ curves lay above or below the $U^{1/2}(\cdot)$ curve seemed to be correlated with the subject's degree of risk aversion, with the $U^{1/2}(\cdot)$ curve typically lying higher relative to the other curves for risk averters and lower for risk lovers. However, since their method of classifying individuals as risk averse or risk loving was based on the concavity or convexity, and hence <u>height</u>, of the $U^{1/2}(\cdot)$ curve, this finding may in part be a statistical artifact introduced by their method of categorizing the observations. Finally, since Allais' method of constructing his "$B_{1/2}$" curves differed slightly from the fractile method, his data may only be used to compare $U^{1/2}(\cdot)$ with $U^{p}(\cdot)$ for $p < 1/2$, where it exhibits the utility evaluation effect described in this section (see Allais (1979, pp. 611-654)).

11. See Note 8.

12. Stochastically dominating shifts in $D\{x_1, x_2, x_3\}$ are shifts which increase p_3 at the expense of p_2 and/or increase p_2 at the expense of p_1, which correspond respectively to upward and/or leftward (i.e., northward and/or westward) shifts in the unit triangle diagram.

13. See Note 6.

REFERENCES

Allais, M., "The Foundations of a Positive Theory of Choice Involving Risk and a Criticism of the Postulates and Axioms of the American School" (Translation of "Fondements d'une Théorie Positive des Choix Comportant un Risque et Critique des Postulats et Axiomes de L'Ecole Americaine', Paris, CNRS), 1952, in Allais & Hagen.

Allais, M., "The So-Called Allais Paradox and Rational Decisions Under Uncertainty," 1979, in Allais & Hagen.

Allais, M., and O. Hagen, Expected Utility Hypotheses and the Allais Paradox: Contemporary Discussions of Decisions Under Uncertainty with Allais' Rejoinder (Dordrecht, Holland: D. Reidell, 1979).

Bolker, E. "A Simultaneous Axiomatization of Utility and Subjective Probability," Philosophy of Science 34 (1967), 333-40.

Chew, S. H., and K. 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, 1979.

Debreu, G., Theory of Value: An Axiomatic Analysis of General Equilibrium (New Haven: Yale University Press, 1959).

Dreze, J., "Axiomatic Theories of Choice, Cardinal Utility and Subjective Probability: A Review." In J. Dreze, ed., Allocation Under Uncertainty: Equilibrium and Optimality (New York: John Wiley and Sons, 1974).

Edwards, W., "The Prediction of Decisions Among Bets," Journal of Experimental Psychology 50 (1955), 201-14.

Farquhar, P., "Utility Assessment Methods," University of California, Davis, Graduate School of Administration Working Paper No. 81-5, 1982.

Fishburn, P., "Nontransitive Measurable Utility," Bell Laboratories Economics Discussion Paper No. 209, Journal of Mathematical Psychology (forthcoming, 1981a).

Fishburn, P., "Transitive Measurable Utility," Bell Laboratories Economics Discussion Paper No. 224, Journal of Economic Theory (forthcoming, 1981b).

Friedman, M. and L. Savage, "The Utility Analysis of Choices Involving Risk," Journal of Political Economy 56 (1948), 279-304. Reprinted in G. Stigler and K. Boulding, eds., Readings in Price Theory (Chicago: Richard D. Irwin, 1952).

Grether, D., "Recent Psychological Studies of Behavior Under Uncertainty," American Economic Review Papers and Proceedings 68 (1978), 70-74.

Grether, D., and C. Plott, "Economic Theory of Choice and the Preference Reversal Phenomenon," American Economic Review 69 (1979), 623-38.

Hadar, J., and W. Russell, "Rules for Ordering Uncertain Prospects," American Economic Review 59 (1969), 25-34.

Handa, J., "Risk, Probabilities, and New Theory of Cardinal Utility," Journal of Political Economy 85 (1977), 97-122.

Hagen, O., "Towards a Positive Theory of Preferences Under Risk," 1979, in Allais & Hagen.

Hersten, I., and J. Milnor, "An Axiomatic Approach to Measurable Utility," Econometrica 21 (1953), 291-97.

Kahneman, D., and A. Tversky, "Prospect Theory: An Analysis of Decision Under Risk," Econometrica 47 (1979), 263-91.

Karmarkar, U., "The Effect of Probabilities on the Subjective Evaluation of Lotteries," Massachusetts Institute of Technology Sloan School of Management Working Paper No. 698-74, 1974.

MacCrimmon, K., "Descriptive and Normative Implications of the Decision Theory Postulates," in K. Borch and J. Mossin, eds., Risk and Uncertainty (London: MacMillan & Co., 1968).

MacCrimmon, K., and S. Larsson, "Utility Theory: Axioms Versus 'Paradoxes,'" 1979, in Allais & Hagen.

Machina, M., "'Rational' Decision Making vs. 'Rational' Decision Modelling?: A Review of Expected Utility Hypotheses and the Allais

Paradox, Edited by Maurice Allais and Ole Hagen," Journal of Mathematical Psychology 24 (1981), 163-75.

Machina, M., "'Expected Utility' Analysis without the Independence Axiom," Econometrica 50 (1982a), 277-323.

Machina, M., "A Stronger Characterization of Declining Risk Aversion," Econometrica 50 (1982b), 1069-79.

Machina, M., "Temporal Risk and the Nature of Induced Preferences," University of California, San Diego, manuscript, 1982c.

Markowitz, H., "The Utility of Wealth," Journal of Political Economy 60 (1952), 151-58.

McCord, M., and R. de Neufville, "Fundamental Deficiencies of Expected Utility Decision Analysis," Massachusetts Institute of Technology, manuscript, 1982.

Morrison, D., "On the Consistency of Preferences in Allais' Paradox," Behavioral Science 12 (1967), 373-83.

Moskowitz, H., "Effects of Problem Representation and Feedback on Rational Behavior in Allais and Morlat-Type Problems," Decision Sciences 5 (1974), 225-42.

Pratt, J., "Risk Aversion in the Small and in the Large," Econometrica 32 (1964), 122-36.

Raiffa, H., Decision Analysis: Introductory Lectures on Choice Under Uncertainty (Reading, MA: Addison-Wesley, 1968).

Rothschild, M., and J. Stiglitz, "Increasing Risk: I. A Definition," Journal of Economic Theory 2 (1970), 225-43.

Savage, L., Foundations of Statistics (New York: Dover Publications, 1972). (Revised and enlarged version of the work originally published by John Wiley & Sons, New York, 1954).

Slovic, P. and A. Tversky, "Who Accepts Savage's Axiom?" Behavioral Science 19 (1974), 386-73.

Tversky, A., "Intransitivity of Preferences," Psychological Review 76 (1969), 31-48.

Tversky, A., "A Critique of Expected Utility Theory: Descriptive and Normative Considerations," Erkenntnis 9 (1975), 163-73.

von Neumann, J., and O. Morgenstern, Theory of Games and Economic Behavior (Princeton: Princeton University Press, 1944). (Second Edition, 1947; Third Edition, 1953).

Wold, H., "Ordinal Preferences or Cardinal Utility?" Econometrica 20 (1952), 661-63.