Journal of Economic Behavior and Organization 3 (1982) 323–343. North-Holland

A THEORY OF ANTICIPATED UTILITY

John QUIGGIN*

Australian Bureau of Agricultural Economics, Canberra City, ACT 2601, Australia

Received July 1980, final version received December 1982

A new theory of cardinal utility, with an associated set of axioms, is presented. It is a generalization of the von Neumann–Morgenstern expected utility theory, which permits the analysis of phenomena associated with the distortion of subjective probability.

1. Introduction

The expected utility theory of von Neumann and Morgenstern (1944) (hereafter NM) is a powerful tool for the analysis of decisions under risk. However, people in both experimental and real-life situations frequently do not conform to the NM axioms. A number of decision problems have been designed for which most people make choices violating the NM axioms. The most celebrated of these is the Allais paradox [Allais (1953)]. Similar problems, which do not involve the extremely large amounts of money used in Allais' problem, have been constructed by MacCrimmon (1968), Kahneman and Tversky (1979), Schoemaker and Kunreuther (1979), and Slovic et al. (1977). Kunreuther et al. (1978) examined the insurance behaviour of home owners in flood-prone and earthquake-prone areas, and found results inconsistent with the NM theory.

These observations are clearly related to the well-known fact that many people who are normally risk-averse are willing to engage in gambles at long odds. Friedman and Savage (1948) offer an ingenious explanation using an S-shaped utility curve, but this raises as many problems as it solves. For example, what about people whose present wealth does not lie in the convex segment of the curve?

One explanation which is intuitively appealing, is that individuals tend to substitute 'decision weights' for probabilities. This has been argued by Fellner (1961) and Edwards (1962). It has also been recognised by practical users of expected utility theory. Thus, Officer and Halter (1968) and

*I would like to thank Ted Sieper, Jock Anderson and two referees for their helpful comments and criticism.

Anderson, Dillon and Hardaker (1977, pp. 69–70) suggest that questionnaires used in eliciting risk preferences should use only 50–50 choices, i.e., gambles with two outcomes each occurring with equal probability.

Unfortunately, it has proved difficult to construct a theory which embodies this insight. Handa (1977) attempted this, but his theory was shown by Fishburn (1978) to imply maximisation of expected returns when violations of dominance were excluded. Karmarkar (1978, 1979) refined Handa's theory with his Subjectively Weighted Utility model, but this reduces to the NM theory when violations of dominance are excluded (see below). The 'prospect theory' of Kahneman and Tversky is more complex and realistic, but it admits 'indirect violations of dominance' and thus intransitivity in pairwise choices.

In this paper, an axiomatic approach to the problem is adopted. It has been observed that the most intractable difficulties with the practical application of NM theory are related to the axiom of irrelevance of independent alternatives. In MacCrimmon's experiments, subjects were given the opportunity to reconsider choices which violated the NM axioms. In most cases, e.g., violations of transitivity, they readily agreed that their original choices had been erroneous, but they were unwilling to alter choices which violated the independence axiom. When arguments based upon this axiom were put to them, their answers indicated that they did not accept the validity of the axiom. A substantial proportion of subjects continued to reject the axiom even after an individual discussion with the experimenter. These results were largely replicated by Slovic and Tversky (1974).

In this paper, a new theory is derived, based on a set of axioms weaker than those of NM, and in particular a weaker form of the independence axiom. Because it is based on a weighted sum of utilities formed using 'decision weights', rather than a mathematical expectation it is termed 'anticipated utility' (AU) theory.

Like expected utility theory, it maintains transitivity in pairwise choices and does not admit violations of dominance.

Two basic criticisms of this approach could be made. First, it could be argued that NM theory is sufficient for all practical purposes. This would be the case if violations of NM theory were merely random erroneous judgements, which were avoided in making important decisions or if, because of self-selection, most important economic decisions will be made by individuals conforming to the NM axioms. (It is worth noting that this latter argument may be used to justify the use of expected wealth maximization rather than expected utility maximization.) However, these arguments are based on the premise that, on consideration, all 'rational' people will accept the NM axioms and seek to make decisions in conformity with them. As McCrimmon's experiments showed, this is not the case, at least for business executives.

Alternatively AU theory could be criticised for not being general enough. McCrimmon's experiments showed a number of violations of axioms shared by AU and NM theory, such as intransitivities [also examined by Tversky (1969)] and violations of dominance [Coombs (1975)]. Other behavioural violations of AU theory, such as context effects, have been examined by Hershey and Schoemaker (1980), and Kahneman and Tversky (1979).

There are a number of answers to this second objection. First, in contrast to the NM independence axiom, transitivity and dominance rules command virtually unanimous assent even from those who sometimes violate them in practice. (Some of Tversky's subjects refused to believe that they could have made such errors.)

Second, at least in the case of transitivity violations, these problems appear to arise mostly in unfamiliar situations. In such situations, decision rules which normally produce rational choices may break down.

Third, transitivity and dominance rules are essential to almost all economic theories of decisions under certainty. Thus, if a theory of decision under uncertainty is to be consistent with any of the large body of economic theory which has already been developed (much of which has substantial empirical support), it must satisfy these rules.

There are, however, a number of generalisations of AU theory which could be considered without violating these rules. Indeed, the anticipated utility function developed here is one of a number of possibilities considered (but not developed or axiomatized) by Allais (1952).

The principal difference between AU theory, and previous analyses of 'decision weights', such as that of Handa, lies in the fact that these approaches sought to derive weights as a function of individual probabilities. This created immediate difficulties with probabilities of the form $1/n$. Severe violations of dominance are generated unless the weight, $w(1/n)$ is set equal to $1/n$. If, say, the weight is less than $1/n$, we may consider an initial situation in which a sum of money, X, is received with certainty, and compare it with the situation where this sum is augmented by a small but positive random variable x, $0<x\leqq\varepsilon$, where x takes n different values x_i each with probability $1/n$. For sufficiently small ε it is apparent that

$$U(X)>\sum_{i=1}^{n} w(1/n)U(X+x_i). \tag{1*}$$

That is, the certain event is preferred, even though the random variable has a superior outcome with probability 1.

If this anomaly is to be avoided, a similar argument will establish the requirement that $w(k/n)=k/n$, for $1\leqq k\leqq n$. Consider the choice between a fixed sum X received with probability k/n (with nothing received otherwise), and k sums of the form $X+x_i$ (or $X-x_i$), each with probability $1/n$. Thus, if violations of dominance are to be avoided, a decision weighting function of this type must be an identity function, and the resulting theory is identical to that of NM. A formal proof of this is given by Fishburn (1978).

AU theory avoids these difficulties, because probability weights are derived from the entire probability distribution, rather than from individual probabilities. Typically, events at extremes of the range of outcomes are likely to be 'overweighted'. In such cases, at least some 'intermediate' outcomes, perhaps with the same objective probability, must be underweighted.

This may be illustrated by the following example. Suppose an individual's normal wage income is uniformly distributed over a range from \$20 000 to \$21 000. There is also a 1/100,000 chance that s/he will win a contest for which the prize is a job paying \$1 000 000 a year. This probability is, in fact, precisely equal to that of receiving any specified income in the relevant range, e.g., \$21 439.72. Nevertheless, we would not expected the two events to be treated in the same way. The factors which might lead someone to overweight the probability of winning the contest, would not be likely to have the same effect on an 'intermediate' outcome, even with the same objective probability. Because two events with the same objective probability need not have the same weight, the problems described above do not apply to AU theory. While this point is fairly easy to grasp intuitively, formalizing it presents some difficulties. These are addressed in the following section.

2. An outline of the theory

The theory deals with individual preferences over a set X of outcomes, and an associated set Y of prospects.[1] The set X of outcomes provides a partition of the possible states of the world into mutually exclusive events. Each prospect $y \in Y$ consists of a pair of vectors $\{(x_1, x_2, \ldots, x_n); (\mathrm{p}_1, \ldots, \mathrm{p}_n)\}$, such that $x_1, x_2, \ldots, x_n \in X$ and $\sum_i p_i = 1$. If the prospect y is selected, the individual will receive outcome x_i with probability p_i. The number of distinct possible outcomes, n, will not, in general, be the same for different prospects.[2] In particular, prospects of the form $\{(x); (1)\}$ in which an outcome is received with certainty, will play an important role. For simplicity, such prospects will not be distinguished from the corresponding elements $x \in X$, and X will be treated as a subset of Y.

Individual preferences are denoted by a relationship P which is assumed to be complete, reflexive and transitive, and by the associated indifference relationship I. If an outcome $c \in X$ is indifferent to a risky prospect $y \in Y$, it

[1]My use of the term 'prospect' differs from that of Kahneman and Tversky who use it to denote a change from some initial position.

[2]Thus, if $X = R$, Y is a subspace of the space of finite-dimensional simplices over R. See Dugundji (1966, ch. 15).

will be called the certainty equivalent of y, denoted $c=CE(y)$. If two outcomes are indifferent they will not be distinguished.

The object of a utility theory is the construction of a function V on Y such that $V(y) \geqq V(y')$ if, and only if, yPy'. The NM theory involves constructing a function U on X and setting $V(y)=\sum_i p_i U(x_i)$. Handa offers the seemingly symmetrical proposal $V(y)=\sum_i w(p_i)x_i$, where the x_i are assumed to represent some quantity of a given good and w is a real-valued function on the unit interval such that $w(0)=0$. Kahneman and Tversky combine the two, setting $V(y)=\sum_i w(p_i)U(x_i)$, where the x_i are assumed to represent changes from some initial situation. Karmarkar suggests setting

$$V(y)=\sum_i w(p_i)U(x_i)\bigg/\sum_i w(p_i), \quad \text{where}$$

$$w(p)=p^\alpha/[p^\alpha+(1-p)^\alpha], \qquad 0<\alpha \leqq 1. \tag{1}$$

Each of these alternatives to NM theory encounter the problems mentioned above. Kahneman and Tversky themselves point to the difficulties which arise with their approach (and *a fortiori* that of Handa) when w is non-linear. Further discussion is given by Karmarkar (1979).

In general, if w is non-linear it is possible to find an 'overweighted' pair of probabilities p and $1-p$ such that

$$w(p)+w(1-p)>1. \tag{2}$$

If x_1 and x_2 are chosen so that x_1 is very slightly preferred to x_2, then

$$U(x_1)<w(p)u(x_1)+w(1-p)U(x_2). \tag{3}$$

Yet the prospect x_1 clearly dominates $\{(x_1,x_2); (p,1-p)\}$ in the strong sense that its outcome will be preferred (or at least indifferent) with probability 1. Kahneman and Tversky avoid this implication by assuming that dominated prospects are 'edited out' but this leads to the undesirable result that pairwise choices are intransitive. Note again that, if w is linear, so that $w(p)+w(1-p)=w(1)$ the theory reduces to that of NM.

A more complex example shows that the problem of dominance applies to Karmarkar's theory. Suppose that, for some p, $w(p)<2w(p/2)$. Then we can find $x_1Px_2Px_3$ such that

$$(w(p)U(x_1)+w(1-p)U(x_3))/(w(p)+w(1-p))$$
$$<(w(p/2)U(x_1)+w(p/2)U(x_2)+w(1-p)U(x_3))/(2w(p/2)+w(1-p)). \tag{4}$$

All that is required is that $U(x_2)$ should be very close to $U(x_1)$ and substantially greater than $U(x_3)$. However, $\{(x_1,x_3);\ (p,1-p)\}$ dominates $\{(x_1,x_2,x_3);(p/2,p/2,1-p)\}$. (Note that this dominance result does not require the NM axiom of independence of irrelevant alternatives. See section 3.) Since a similar result applies if $w(p)>2w(p/2)$, preservation of dominance implies

$$w(p_i)\Big/\sum_i w(p_i)=p_i \quad \text{for all} \quad p=(p_1,p_2,\ldots,p_n). \tag{5}$$

In this case, Karmarkar's approach is identical to that of NM.

All of these counter-examples require some assumptions on the 'richness' of the set X of outcomes, such as those made by Fishburn (1978). However, they will always apply if X is an interval on the real line, and any viable theory must be able to cover this case.

As was stated above, the fundamental problem in these theories is that any two outcomes with the same probability must have the same decision weight. This fails to take account of the fact that, while individuals may distort the probability of an extreme outcome in some way, they need not treat 'intermediate' outcomes with the same probability in the same fashion.

In order to formalize this observation, it is necessary to order the possible outcomes x_i, and the corresponding probabilities, p_i, in each prospect. We assume that the outcomes are ordered from worst to best, i.e., $x=(x_1,x_2,\ldots,x_n)$, where $x_nPx_{n-1}P,\ldots,Px_2Px_1$ and $p=(p_1,p_2,\ldots,p_n)$, where p_i is the probability of outcome x_i.

The anticipated utility function is defined to be

$$V=h(p)\cdot U(x)=\sum_i h_i(p)U(x_i), \tag{6}$$

where U is a utility function with properties similar to that of NM, while $h(p)$ is a vector of decision weights satisfying $\sum_i h_i(p)=1$. In general, $h_i(p)$ depends on all the p_js and not just on p_i. Thus, for example, the fact that $p_j=p_k$ would not imply that $h_j(p)=h_k(p)$. It is obviously necessary to require that if $p_i=0$, then $h_i=0$. For simplicity of scaling, it will be assumed that $h(1)=1$.

More significantly, it will be assumed that $h(\frac{1}{2},\frac{1}{2})=(\frac{1}{2},\frac{1}{2})$. The claim that the probabilities of 50–50 bets will not be subjectively distorted seems reasonable, and, as stated above, has proved a satisfactory basis for practical work [Anderson, Dillon and Hardaker (1977)]. The experiments discussed by Handa yielded crossover points such that $h_1(p)=(p_1$ for values of p_1 less than $\frac{1}{2}$ but they did not take risk aversion into account. As discussed below, risk aversion can easily be confused with pessimism [roughly, setting $h_i(p)<p_i$ for the more favourable outcomes].

The description of the function h would at first sight appear to be an infinite task since no bound has been imposed on the number of different outcomes, n, in a given prospect, y. However, it can be shown that a knowledge of $h(p, 1-p)$ for each $p \in [0,1]$ is sufficient to determine $h(p)$ for any probability vector p, regardless of its length. The approach used is very similar to that which was used to show that a non-linear weighting function on individual probabilities will generate violations of dominance. The central point is that the evaluation of a prospect, in which two very similar outcomes x_1 and x_2 occur with the probabilities p_1 and p_2, must be 'close' to that of a prospect which is identical except that x_1 replaces x_2, occurring with a total probability p_1+p_2. This imposes constraints on the function h. If it is also assumed, as in previous approaches, that $h_i(p)$ depends only on p_i, then these constraints are satisfied only by the NM expected utility function.

For each $p \in [0,1]$, write $f(p)=h_1(p,1-p)$. Thus, $f(p)$ defines the behaviour of h on pairs $(p, 1-p)$. The extension of h to triples (p_1, p_2, p_3) will now be described in detail. The approach used can be shown by induction to apply to arbitrary p.

Consider the two prospects $y_1=\{(x_1,x_2);\ (p_1,1-p_1)\}$ and $y_2=\{(x_1,x_3,x_2);\ (p_2,p_1-p_2,1-p_1)\}$. As $U(x_3)$ approaches $U(x_1)$, $V(y_2)$ must approach $V(y_1)$. At the limit, $x_1=x_3$ and $y_1=y_2$, so that

$$\begin{aligned} f(p_1)U(x_1)+(1-f(p_1))U(x_2) &= (h_1(p_2,p_1-p_2,1-p_1) \\ &\quad + h_2(p_2,p-p_2,1-p_1))U(x_1)+h_3(p_2,p_1-P_2,1-p_1)U(x_2). \end{aligned} \tag{7}$$

Since x_1 and x_2 are chosen arbitrarily and $h(p)$ is independent of x the coefficients on $U(x_2)$ on the right-hand and left-hand sides of (7) must be equal. Hence,

$$h_3(p_2,p_1-p_2,1-p_1)=1-f(p_1). \tag{8}$$

Conversely, by setting $U(x_3)$ close to $U(x_2)$ we can show

$$h_1(p_2,p_1-p_2,1-p_1)=f(p_2), \quad \text{and hence} \tag{9}$$

$$h(p_1,p_2,p_3)=(f(p_1), f(p_1+p_2)-f(p_1), 1-f(p_1+p_2)).$$

More generally:

$$h_i(p)=f\left(\sum_{j=1}^{i} p_j\right)-f\left(\sum_{j=1}^{i-1} p_j\right) \tag{10}$$

so that the behaviour of h on arbitrary probability distributions is fully determined by the values of $h(p, 1-p)$ for $0<p<1$.

The result shows that $h_i(p)$ depends on all the probabilities $p_1, p_2, \ldots, p_n$ and not merely p_i. Moreover the relationship between the probability of an event x_i and its decision weight depends upon its position in the preference ranking of possible outcomes.

Eq. (10) also offers a natural extension to the case where Y includes continuous probability distributions, a case which cannot be handled at all in the theories of Handa and Kahneman–Tversky.

Let $D(x_0)=P_r\{U(x)<U(x_0)\}$ be the (objective) cumulative distribution function for x. Then D ranges from 0 to 1, and so the function f may be applied to obtain a 'cumulative distribution function' for decision weights, G. Thus

$$G(x_0)=f(D(x_0). \tag{11}$$

Let the density function associated with G be written g. This function plays the same role for continuous probability distributions as h plays for discrete ones. Thus, anticipated utility is given by

$$V=\int_x^{\infty} U(x)g(x)\,\mathrm{d}x=\int_x^{\infty} U(x)\,\mathrm{d}G(x). \tag{12}$$

This result can be proved easily if the integral in (12) is a Riemann integral over a bounded interval $[a,b]$. This is because the integral is formed as a limit (from above and below) of step function integrals which may be related to elements of Y with discrete distributions. In fact we may construct series y_N^* and y_N^{**} such that $y_N^* P y P y_N^{**}$ for all N and $\lim_{N\to\infty} V(y_N^*=\lim_{N\to\infty} V(y_N^{**})=V$ so that $V(y)=V$ as required.

Assume without loss of generality that $[a,b]=[0,1]$ and partition the interval into N sub-intervals, so that

$$I_{iN}=[(i-1)/N, i/N].$$

The probability that y will lie in I_{iN} is given by

$$\begin{aligned} P_{iN}&=\Pr\{y\in I_{iN}\}\\ &=D(i/N)-D((i-1)/N) \qquad i=1,2,\ldots,N. \end{aligned} \tag{13}$$

By (11),

$$\begin{aligned} h_i(\boldsymbol{p}_N)&=f(D(i/N))-f(D(i-1)/N))\\ &=g(x_{iN}^*)/N \quad \text{for some } x_{iN}^*\in I_{iN} \end{aligned} \tag{14}$$

by the Mean Value Theorem. Define $y_N^{**}=(0,\ldots,N-1/N;p_N)$ and $y_N^*=(i/N,\ldots,1;p_N)$.

The distributions of y_N^* and y_N^{**} are given by step functions which lie, respectively, above and below D. It is clear that $y_N^* Py Py_N^{**}$, while

$$V(y_N^{**})=\sum_{i=1}^{N} h_i(p_N)U(i/N)$$

$$=\frac{1}{N}\sum_{i=1}^{N} g(x_{iN}^*)U(i/N), \quad \text{and} \tag{15}$$

$$V(y_N^{**})=1/N\sum_{i=1}^{N} g(x_{iN}^*)U((i-1)/N), \quad \text{so that}$$

$$\lim_{N\to\infty} V(y_N^*)=\lim_{N\to\infty} V(y_N^*)=V.$$

3. Axioms

In order to generate a theory more general than NM expected utility theory, it is necessary to begin with a weaker set of axioms. In particular, it seems appropriate to modify the controversial independence axiom.

The notation adopted here creates some difficulties in comparisons of the two sets of axioms. First, it must be noted that acceptance of the NM complexity axiom (that the value of a compound lottery depends only on the ultimate probability of each outcome) is implicit in the use of this notation. If this axiom did not hold, two prospects, both expressed as (x,p), need not be of equivalent value. Second, the independence and continuity axioms in NM theory are significantly weakened if they apply only to outcomes $x\in X$.

For instance, the independence[3] axiom:

$$y_1Py_2 \Rightarrow y_1P\{(y_1,p),(y_2,1-p)\} \quad \text{for any } p\in[0,1].$$

becomes a simple statement that preferences preserve dominance if y_1 and y_2 are replaced by elements of X (see Axiom 2).

In addition to the complexity axiom, the usual completeness axiom will be adopted.

Axiom 1. P is complete, reflexive and transitive (completeness).

[3]A widely used version of the independence axiom is: $Y_1(Py_2 \Rightarrow \{(y_1,p),(y_3,1-p)\}P\{y_2,p),(y_3,1-p\}$ for any $y_3\in Y$ and $p\in[0,1]$. It is this version which is directly contradicted by Allais' results. In combination with the other NM axioms, this version of the independence axiom is equivalent to that given above.

The dominance and continuity axioms will be modified to apply to outcomes only.

Axiom 2. $x_1, x_2 \in X, x_1 P x_2 \Rightarrow x_1 P\{(x_2, x_1); (p, 1-p)\}$ (dominance).

Axiom 3. If $x_1, x_2, x_3 \in X, x_1 P x_2 P x_3$ there exists p^*, such that

$$x_2 I\{(x_3, x_1); (p^*, 1-p^*)\} \qquad \text{(continuity)}.$$

Finally, a weak independence axiom will be used to ensure that $h(p)$ is independent of x.

Axiom 4. If $y_1 = \{x, p\}$ and $y_2 = \{x', p\}$ and for each $i, i = 1, 2, \ldots, n$, there exists

$$c_i = CE\{x_i, x_i'); (\tfrac{1}{2}, \tfrac{1}{2})\}, \qquad i = 1, 2, \ldots, n, \quad \text{and}$$

$$x_1^* = CE(y_1), \qquad x_2^* = CE(y_2). \quad \text{Then}$$

$$\{c; p\} I\{(x_1^*, x_2^*); (\tfrac{1}{2}, \tfrac{1}{2})\} \qquad \text{(independence)}.$$

Thus, if each element of c is indifferent to a 50–50 bet consisting of corresponding elements of x and x', $\{c;p)$ is indifferent to a 50–50 bet consisting of certainty equivalents of $\{x;p\}$ and $\{x';p\}$.

Thus far, no assumptions have been made about the 'richness' of X and Y. If they have insufficiently many elements, utility functions may be non-unique and the proof of existence theorems is made more complicated. (An existence theorem can be proved in some cases by inventing additional elements of X and Y and extending P to cover them. A utility function which preserves the extended preference ordering will also preserve the original ones.)

The following assumptions will be made:

R.1. If $x_1, x_2, \ldots, x_n \in X$, $x_n P x_{n-1} P, \ldots, P x_1$, and $\sum_i p_i = 1$, then $\{x; p\} \in Y$.

R.2. If $y \varepsilon Y$ there exists $x = CE(y) \in X$.

These assumptions are stronger than those required to derive uniqueness results in NM theory because the class of utility functions to be considered is much larger. However, some form of R.1 is required for both theories and R.2 will always be satisfied for the standard situation where X is a connected subset of the real line.

The major result of this paper is:

Proposition 1. Suppose X and Y satisfy R.1 and R.2 and P satisfies Axioms 1–4. Then there exists a function $V: Y \to R$ *such that*

(i) $V(y) \geqq V(y')$ *if and only if* yPy',

(ii) $V(x;p) = \sum_i h_i(p)U(x_i)$ *for some functions U and h,*

where $h(1)=1$ *and* $h(\frac{1}{2},\frac{1}{2}) = (\frac{1}{2},\frac{1}{2})$. *If two functions V and V′ satisfy (i) and (ii) there exists constants a and b,* $a>0$, *such that*

(iii) $V'(y) = aV(y) + b$.

The proof, which is not very illuminating, is included in the appendix.

It is clear that Axioms 1–3 are weak forms of the corresponding NM axioms. Axiom 4 can be derived from the NM independence axiom but it is easier to show that any expected utility maximizer must satisfy it, since

$$E[U\{(X_1^*, X_2^*); (\tfrac{1}{2},\tfrac{1}{2})\}] = \tfrac{1}{2}\sum_i U(x_i)p_i + \tfrac{1}{2}\sum_i U(x_i')p_i$$

$$= \sum_i \tfrac{1}{2}(U(x_i) + U(x_i'))p_i = E[U\{c;p\}]. \tag{16}$$

4. Conditions on the anticipated utility functions

Under the anticipated utility theory, an individual's attitudes to prospects are determined both by their attitudes to the possible outcome and by their attitudes to the probabilities. These are reflected in the utility function, u, and the weighting function, h, respectively.

With regard to the first, the concepts of risk aversion, risk neutrality and risk preference are still relevant though their interpretation is somewhat different. A risk neutral individual will be indifferent between a 50–50 bet and its expected outcome, while a risk averter will prefer the certain outcome and a risk preferrer the bet. Thus, as in expected utility theory, risk aversion is equivalent to a concave utility function.

The most obvious pattern of probability distortion relates to the treatment of events with small probabilities and extreme outcomes. We may say that an individual overweights extreme events if

$$\text{(A)} \quad \begin{cases} f(p) \geqq p, & p \leqq \tfrac{1}{2}, \\ f(p) \leqq p, & p \geqq \tfrac{1}{2}. \end{cases}$$

Conversely, an individual underweights extreme events if

$$\text{(B)}\qquad \begin{cases} f(p)\leqq p, & p\leqq\tfrac{1}{2}, \\ f(p)\geqq p, & p\geqq\tfrac{1}{2}. \end{cases}$$

Of course, an individual may conform to neither of these patterns. For example, there may be a number of probabilities for which $f(p)=p$. Even if (A) or (B) is satisfied there may be multiple points of inflexion in addition to that at $(\frac{1}{2},\frac{1}{2})$.

To avoid the latter possibility the stronger assumptions

(A*) f is concave on $[0,\frac{1}{2}]$ and convex on $[\frac{1}{2},1]$, and

(B*) f is convex on $[0,\frac{1}{2}]$ and concave on $[\frac{1}{2},1]$

may be used in place of (A) and (B) respectively.

In general, overweighting of extreme events would seem to be the norm. Overweighting of extreme events provides a simple explanation for the Friedman–Savage and Allais paradoxes. On the other hand, it often seems as if events with extremely small probabilities (e.g., less than 10^{-5}) are simply ignored. Kahneman and Tversky discuss this phenomenon, which they describe as 'editing', and suggest it implies that f must be discontinuous near 0 and 1.

I wish to offer a reinterpretation of the 'editing' process which explains both overweighting of extreme events with small probabilities and disregarding those of extremely small ones. It would appear that there is a non-zero-cost of calculation which is incurred if an event is included in our calculations. Hence, events which affect anticipated utility by an amount less than this cost are edited out. Such events either are extremely improbable or have only a miniscule effect on ulity if they do occur.

However, intuitive processes, such as editing, will be refined by trial and error, rather than by deductive reasoning. For 'once in a lifetime' events, such as winning the lottery, this process will not be very effective. There is likely to be a 'conservative' bias against excluding such events from consideration, unless they are very improbable indeed. In the case of lotteries, of course, the editing process will also be counteracting by advertising, which seeks to keep the possibility of winning always in the minds of customers.

Thus, it would appear, the editing process will eliminate some improbable events with extreme outcomes, but, from an objective probability viewpoint, not enough.

The fact that events with extremely low probabilities are ignored does not mean f must be discontinuous. If the range of utility outcomes is bounded, all events with a sufficiently low probability will have a minimal effect on anticipated utility, but it is not necessary to set $f=0$ to represent this.

A second feature of the transformation f is optimism-pessimism. If, for all p, $f(p) \geqq 1-f(1-p)$ an individual is pessimistic, since the worst outcomes are, on average, overweighted. If $f(p) \leqq 1-f(1-p)$, the individual is optimistic, while if $f(p)=1-f(1-p)$ the individual is neutral and h is symmetric.

As mentioned above, pessimism is rather difficult to distinguish empirically from risk aversion. It would thus be possible to require U to be linear, as Handa did. However, this would imply dropping the plausible condition $h(\frac{1}{2},\frac{1}{2})=(\frac{1}{2},\frac{1}{2})$ and would also imply constant relative risk aversion for any given odds. The theory presented here is more general and flexible.

5. Stochastic dominance

Most useful applications of the NM theory have been based on the assumption that X is the set of real numbers. (Its elements are generally taken to correspond to levels of real wealth.) The preference ordering P on X is assumed to correspond to the natural ordering, so that individuals always prefer more wealth to less. These preferences may be represented by a monotonic increasing function $U: R \rightarrow R$. One of the most important tools of analysis under these conditions is the concept of stochastic dominance [Fishburn (1964)]. Hadar and Russell (1969) have developed stochastic dominance rules for the NM theory. Given random variables y_0 and y_1, with cumulative distribution functions D_0 and D_1 respectively, they define y_1 to first stochastically dominate y_0 (y_1 FSD y_0) if

$$D_0(x) \geqq D_1(x) \quad \text{for all } x \tag{17}$$

and to second stochastically dominate y_0 (y_1 SSD y_0) if

$$\int_{-\infty}^{x} D_0(s)\,\mathrm{d}s \geqq \int_{-\infty}^{x} D_1(s)\,\mathrm{d}s \quad \text{for all } x. \tag{18}$$

(The terms first degree and second degree stochastic dominance are also widely used for these conditions.)
They prove[4]

Proposition 2. Let y_1 and y_0 be as above. Then

(i) *y_1 FSD y_0 if and only if $E[U(y^1)] \geqq E[U(y_0)]$ for all monotonic increasing U.*

(ii) *y_1 SSD y_0 if and only if $E[U(y_1)] \geqq E[U(y_0)]$ for all concave increasing U.*

[4]Some technical errors in their proof are corrected by Tesfatsion (1976).

Some results of this type can be proved in the AU theory. The FSD result extends simply.[5]

Proposition 3. *Let y_0 and y_1 be as above. Then y_1 FSD y_0 if and only if $V(y_1) \geqq V(y_0)$ for all anticipated utility functions V.*

Proof. 'If'. For any U the function $V(y) = E[U(y)]$ is an anticipated utility function and thus Proposition 2(i) applies.

'Only if'. Suppose y_1 FSD y_0. Then for any transformation f

$$G_0(x) = f(D_0(x)) \geqq f(D_1(x)) = G_1(x) \quad \text{for all } x. \tag{19}$$

Thus by Proposition 2(i) $\int U(x)\,\mathrm{d}G_1(x) \geqq \int U(x)\,\mathrm{d}G_0(x)$ for any monotonic increasing U. Thus by (11), $V(y_1) \geqq V(y_0)$ for any anticipated utility function V.

The SSD case is not so simple. Consider the class V^* of AU functions formed from a concave utility function U and a symmetric transformation h which satisfies condition (A^*). If y_0 is a lottery ticket with a large prize, and y_1 is its actuarial expected value then y_1 SSD y_0, but for some $V \in V^*$, $V(y_0) > V(y_1)$. Thus a complete extension of 2(ii) is not possible (or desirable, since the object of the AU theory is the analysis of phenomena such as this). Some partial extensions may, however, be obtained.

The basic approach is contained in the following useful special case:

Proposition 4. *Let y_0 and y_1 be random variables with equal medians and symmetric distributions and with distribution functions $F_0(s)$ and $F_1(s)$ which intersect finitely often. Then y_1 SSD y_0 if and only if $V(y_1) \geqq v(y_0)$ for each $V \in V^*$.*

Proof. Since y_0 and y_1 are symmetric,

$$E[y_0] = \text{median } (y_0) = \text{median } (y_1) = E[y_1] = \mu. \tag{20}$$

Let y_1 SSD y_0. For any $x \leqq \mu$ we may define a sequence $-\infty = a_0 \leqq a_1 \leqq \cdots \leqq a_n = x$ such that

$$\begin{aligned} &s \in (a_{2i}, a_{2i+1}) \Rightarrow D_0(s) \geqq D_1(s)), \qquad i = 1, 2, \ldots, \\ &s \in (a_{2i+1}, a_{2i+2}) \Rightarrow D_1(s) \geqq D_0(s)). \end{aligned} \tag{21}$$

[5]Analogous definitions and results can be given for discrete probability distributions. Since the proofs are also analogous, only the continuous case will be examined here.

It will now be shown that for any f concave on $[0,\frac{1}{2}]$,

$$\int_{a_0}^{a_{2i}} f(D_0(s))-f(D_1(s))\,\mathrm{d}s \geqq f'(D_0(a_{2i-1}))\int_{a_0}^{a_{2i}} (D_0(s)-D_1(s))\,\mathrm{d}s. \tag{22}$$

Let $s\in(a_0,a_1)$. Then by the mean value theorem,

$$f(D_0(s)-f(D_1(s))=f'(c)(D_0(s)-D_1(s)) \tag{23}$$

for some c such that $D_1(s)\leqq c\leqq D_0(s)\leqq D_0(a_1)$. By the concavity of f, $f'(c)\geqq f'(D_0(a_1))$ and

$$\int_{a_0}^{a_1} f(D_0(s))-f(D_1(s))\,\mathrm{d}s \geqq f'(D_0(a_1))\int_{a_0}^{a_1} (D_0(s)-D_1(s))\,\mathrm{d}s. \tag{24}$$

By a converse argument, for $s\in(a_1,a_2)$,

$$f(D_1(s))-f(D_0(s))\leqq f'(D_0(a_1))(D_1(s)-D_0(s)), \quad \text{and} \tag{25}$$

$$\int_{a_1}^{a_2} f(D_0(s))-f(D_1(s))\,\mathrm{d}s \geqq f'(D_0(a_1))\int_{a_1}^{a_2} D_0(s)-D_1(s)\,\mathrm{d}s. \tag{26}$$

Combining (23) and (25) yields the desired result (22) for $i=1$. Suppose (22) holds for $i=1,2,\dots,k$. Then

$$\int_{a_0}^{a_{2k}} (f(D_0(s))-f(D_1(s)))\,\mathrm{d}s \geqq f'(D_0(a_{2k-1}))\int_{a_0}^{a_{2k}} (D_0(s)-{}_1(s))\,\mathrm{d}s. \tag{27}$$

Application of an argument similar to that for $i=1$ to the interval (a_{2k},a_{2k+2}) yields

$$\int_{a_{2k}}^{a_{2k+2}} f(D_0(s)-f(D_1(s))\,\mathrm{d}s \geqq f'(D_0(a_{2k+1}))\int_{a_{2k}}^{a_{2k+2}} (D_0(s)-{}_1(s))\,\mathrm{d}s. \tag{28}$$

Since $f'(D_0(a_{2k}))\geqq f'(D_0(a_{2k+1}))$ we may combine (27) and (28) to yield the desired result for $i=k+1$ and hence by induction, for all i.

Now, if n is even, this shows that

$$\int_{-\infty}^{x} f(D_0(s))-f(D_1(s))\,\mathrm{d}s \geqq 0. \tag{29}$$

If n is odd then (22) holds for $2i=a_{n-1}$ and $f(D_0(s))\geqq f(D_1(s))$ for $s\in(a_{n-1},a_n)$ so that (29) holds in this case also.

Now the symmetry of y_0 and y_1 implies that

$$\int_{\mu}^{\mu+d} (1/2-D_j(s))\,\mathrm{d}s = \int_{\mu-a}^{\mu} (D_j(s)-\tfrac{1}{2})\,\mathrm{d}s, \qquad j=0,1, \tag{30}$$

$$\int_{\mu}^{\mu+a} (\tfrac{1}{2}-f(D_j(s))\,\mathrm{d}s = \int_{\mu-a}^{\mu} (f(D_j(s))-\tfrac{1}{2})\,\mathrm{d}s, \qquad j=0,1 \tag{31}$$

and hence

$$\int_{\mu}^{\mu+a} (f(D_0(s))-f(D_1(s)))\,\mathrm{d}s = -\int_{\mu-a}^{\mu} (f(D_0(s))-f(D_1(s)))\,\mathrm{d}s, \tag{32}$$

i.e.,

$$\int_{-\infty}^{\mu+a} (f(D_0(s))-f(D_1(s)))\,\mathrm{d}s = -\int_{-\infty}^{\mu-a} (f(D_0(s))-f(D_1(s)))\,\mathrm{d}s \geqq 0. \tag{33}$$

The argument of Proposition 3 may now be repeated to yield the desired result.

A useful implication of this result is that, as in the EU theory, mean-variance analysis is valid for comparing normally distributed random variables. For if y_1 has a higher mean and lower variance than y_0 Proposition 3 may be applied to show that $V(y_1) \geqq V(y_0)$ for any V. If the means are equal and y_1 has a lower variance, Proposition 4 may be applied for any $V \in V^*$. Thus indifference curves in the $\mu-\sigma$ plane will be correctly shaped for any $V \in V^*$.

The symmetry assumption may be relaxed considerably. Suppose y_0 is symmetric but F_1 satisfies

$$D_1(\mu+a)-\tfrac{1}{2} \leqq \tfrac{1}{2}-D_1(\mu-a) \quad \text{for all } a>0. \tag{34}$$

By the symmetry of f this implies

$$f(D_1(\mu+a))-\tfrac{1}{2} \leqq \tfrac{1}{2}-f(D_1(\mu-a)), \quad \text{and hence} \tag{35}$$

$$\int_{\mu}^{\mu+a} (f(D_1(s))-\tfrac{1}{2})\,\mathrm{d}s \leqq \int_{\mu-a}^{\mu} (\tfrac{1}{2}-f(D_1(s))\,\mathrm{d}s. \tag{36}$$

Thus (31) may be replaced by

$$\int_{-\infty}^{\mu+a} (f(D_0(s))-f(D_1(s))\,\mathrm{d}s \geqq \int_{-\infty}^{\mu-a} (f(D_0(s))-f(D_1(s)))\,\mathrm{d}s \geqq 0. \tag{37}$$

Condition (34) states that the distribution of y_1 is skewed to the right. [There are distributions with a positive third moment which do not satisfy

(34). However these do not fit the intuitive conception of skewness as well as distributions satisfying (34).] A converse argument can be applied if y_0 is skewed to the left, and y_i is either symmetric or skewed to the right.

6. Concluding comments

Just as the expected utility theory permits the analysis of behaviour which would be excluded as irrational under a profit-maximization hypothesis, the AU theory permits the analysis of anticipations which are not mathematical expectations.

It may be argued that the adoption of a more general theory makes it harder to obtain useful predictions. The results of section 4 show that much of the dominance analysis which is useful in NM theory can be extended to AU theory. In particular, Proposition 3 shows that no pairwise preference ranking which is inadmissible under NM theory can be admissible under AU theory. Further work is needed in developing conditions under which other results can be extended.

Areas in which AU theory might usefully be applied include the problem of individual's apparent propensity to 'over-insure', and the economic analysis of gambling behavior. The theory is likely to be particularly valuable in the analysis of decisions involving catastrophic (or extremely favourable outcomes) which occur with low probability.

Appendix

The proof of Proposition 1 has two parts. In part (a) an arbitrary choice of values for U for two elements of X is shown to imply that $V(y)$ can have only one possible value of any $y \in Y$. In part (b) it is shown that the function V constructed in this way satisfies conditions (i) and (ii) of the proposition. In both parts of the proof the special status of 50–50 bets is used extensively.

(*a*) Choose x', $x'' \in X$, $x''Px'$ and set $U(x')=0$, $U(x'')=1$. [In the trivial case when $x''Ix'$ for all x', x'', V must be a constant function and will satisfy conditions (i) and (ii).]

If $x=XE\{(x',x'');(\frac{1}{2},\frac{1}{2})\}$ then by (ii) and Axiom 2, $U(x)=\frac{1}{2}$. More generally, *if* $U(x_1)=a/2^n$, $U(x_2)=(a+1)/2^n$, $x=CE\{(x_1,x_2);(\frac{1}{2},\frac{1}{2})\}$, $U(x)=(2a+1)/2^{n+1}$. Note that x_2PxPx_1 so that the function U preserves the preference ordering. The set of numbers of the form $a/2^k$ is dense in $[0,1]$ since every real number has a binary expansion. We may thus complete the construction of U on $\{x;x''PxPx'\}$ as follows: Let

$$X_k^*=\{x\in X; U(x)=a/2^k, a=0,1,\dots,2^k\}, \quad \text{and} \tag{A.1}$$

$$X^*=\bigcup_{k=0}^{\infty} X_k^* \text{ (observe } X_k^*<x_{k+1}^*\text{, all } k\text{).} \quad \text{Then} \tag{A.2}$$

$$U(x)=\sup\{U(x_0): x_0 \in X^*, xPx_0\} = \inf\{U(x_0): x_0 \in X^*, x_0Px\}.$$

Note once more that U preserves the preference ordering. Now for any x such that $x''PxPx'$, Axiom 3 requires the existence of $p \in [0,1]$ such that $x = CE\{(x', x''); (p, 1-p)\}$.

Hence, we can determine $f(p)$ by

$$f(1-p) = 1 - f(p) = U(x). \tag{A.3}$$

By the dominance axiom f is monotone increasing. Hence, letting $U(x)$ range from 0 to 1, $f(p)$ can be determined for all p. As shown above, this is sufficient to determine $h(p)$ for all p.

Construction of U outside the range [0, 1], is done using Axiom 3. If xPx'' [so that $U(x) \geqq 1$], there exists p^* such that $x'' = CE\{(x', x; p^*, 1-p^*)\}$. Hence

$$U(x'') = f(p^*)U(x') + (1 - f(p^*))U(x), \quad \text{i.e.,}$$

$$U(x) = 1/(1 - f(p^*)) \quad [\text{since } U(x'') = 1, U(x') = 0]. \tag{A.4}$$

Similarly if $x'Px$, there exists p^* such that

$$x' = CE\{(x, x''); p^*, 1-p^*)\}, \quad \text{and}$$

$$U(x) = -(1 - f(p^*))/f(p^*) = 1 - 1/f(p^*). \tag{A.5}$$

Now U has been determined for all $x \in X$, and h for all $p \in [0,1]$. Hence, given the initial choices $U(x')=0$ and $U(x'')=1$, only one choice is possible for the functions U and h. If, instead, we chose $U(x')=b$, $U(x'')=a+b$ and thus determined a function V', an examination of the construction procedure would show that for all $y \in Y$, $V'(y) = aV(y) + b$. This completes the proof of uniqueness.

(*b*) It has already been shown that U preserves the preference ordering for $x_1, x_2 \in X$ such that $x''Px_2Px'$, $x''Px_1Px'$. But the choice of x'' and x' does not affect the ordering of U [by part (a)] so we can always make this choice to ensure that U preserves the preference ordering for any $x_1, x_2 \in X$.

Similarly for arbitrary $y = \{x, p\}$ it will be shown that a function V based on a selection of x'', x' such that $x'' > x_i \geq x'$ for all i preserves the ordering P. This result must therefore apply to any V.

By transitivity (Axiom 1) and the fact that U preserves P on X, it is sufficient to prove that for any $x \in X$, $y \in Y$, $V(x) = V(y) \Leftrightarrow x = CE(y)$.

Define

$$Y_k^* = \{(x,p\} \in Y: x_i \in X_k^*, \text{all } i\}, \quad \text{and} \tag{A.6}$$

$$Y^* = \{\{x,p\} \in Y: x_i \in X^* \text{ all } i\}. \tag{A.7}$$

The density properties of X^* mean that it will be sufficient to prove the desired result for $y \in Y^*$

The result will first be established for the class of 50–50 bets. Suppose x, x_1, $x_2 \in X_K^*$ and $x_2 P x_2$. Then we wish to prove

Lemma 1. $V(x) = V\{(x_1, x_2); (\frac{1}{2}, \frac{1}{2})\}$ *if and only if* $x = CE\{(x_1, x_2); (\frac{1}{2}, \frac{1}{2})\}$.

Proof. We may note that it is sufficient to prove the 'only if' part of the proposition since, for any x_1, $x_2 \in X_k^*$, there exists x_3 such that $U(x_3) = U\{(x_1, x_2); (\frac{1}{2}, \frac{1}{2})\}$. This can be seen by examining the construction procedure. Hence the 'only if' proposition implies that $x_3 = CE\{(x_1, x_2); (\frac{1}{2}, \frac{1}{2})\}$ and that if $x = CE\{(x_1, x_2); (\frac{1}{2}, \frac{1}{2})\}$ then $U(x) = U(x_3)$. Since $x, x_1 \in X_k^*$,

$$U(x) = a/2^k \qquad \text{for some } a, \quad \text{and}$$

$$U(x_1) = (a-b)/2^k \quad \text{for some } b > 0.$$

The proposition holds for $b=1$ by virtue of the construction procedure. It may be proved inductively for arbitrary b by showing that

(i) if the proposition holds for $b=n$, it holds for $b=2n$,
(ii) if the proposition holds for $b=b_1$ and $b=b_2$ it holds for $b=(b_1+b_2)/2$, when this is an integer.

Any positive integer may be formed using (i) and (ii).

(i) Assume the proposition true for $b=n$ and let $U(x_1) = (a-2n)/2^k$, $U(x_2) = (a+2n)/2^k$.

Define

$$c_1 = CE\{(x_1, x); \tfrac{1}{2}, \tfrac{1}{2})\}, \qquad c_2 = CE\{(x, x_2); (\tfrac{1}{2}, \tfrac{1}{2})\}.$$

Then by the inductive hypothesis,

$$V(c_1) = (a-n)/2^k, \qquad V(c_2) = (a+n)/2^k, \text{ and}$$

$$x = CE\{(c_1, c_2); (\tfrac{1}{2}, \tfrac{1}{2})\}.$$

Application of Axiom 4 with $y_1 = x = \{(x,x);(\frac{1}{2},\frac{1}{2})\}$, $y_2 = \{(x_1,x_2);(\frac{1}{2},\frac{1}{2})\}$ and $x^* = CE(y_2)$ yields xIx^*, that is

$$x = CE\{(x_1,x_2);(\tfrac{1}{2},\tfrac{1}{2})\}, \tag{A.8}$$

which completes part (i),

(ii) Let $b=(b_1+b_2)2$ and

$$U(x_1)=(a-b)/2^k, \quad U(x_2)=(a+b)/2^k,$$

$$U(x_3)=(a-b_1)/2^k, \quad U(x_4)=(a+b_1)/2^k,$$

$$U(x_5)=(a-b_2)/2^k, \quad U(x_6)=(a+b_2)/2^k.$$

Define

$$y_1=\{(x_3,x_5);(\tfrac{1}{2},\tfrac{1}{2})\}, \qquad y_2=\{(x_4,x_6);(\tfrac{1}{2},\tfrac{1}{2})\}. \quad \text{Then}$$

$$x_1=CE(y_1), \qquad x_2=CE(y_2),$$

$$x=CE\{(x_3,x_4);(\tfrac{1}{2},\tfrac{1}{2})\}=CE\{(x_5,x_6);(\tfrac{1}{2},\tfrac{1}{2})\}.$$

Application of Axiom 4 to y_1 and y_2 yields

$$x=CE\{(x_1,x_2);(\tfrac{1}{2},\tfrac{1}{2})\}. \tag{A.9}$$

and the proof of the Lemma is complete.

For the general result for $y \in Y^*$, we use an inductive argument on the Y_K. For $k=0$, the construction of h guarantees the result. Assume the result holds for $k=0, 1,\ldots,n-1$. Let $y=\{x,p\} \in Y_n^*$ and suppose $x_0=CE(y)$. Define x^* and x^{**} as follows: If $x_i \in X_{n-1}^*$, $x_1^*=x_i=x_i^{**}$. If $x_i \in X_n^*-X_{n-1}^*, U(x)=(2a+1)/2^n$ for some a. Choose x_i^*, x_i^{**} such that $U(x_i^*)=a/2^{n-1}$, $U(x_i^{**})=(a+1)/2^{n-1}$. Then $\{x^*,p\}$ and $\{x^{**},p\}$ are members of Y_{n-1}^*. They also have the properties that

(i) $V\{x,p\}=\frac{1}{2}V\{x^*,p\}+\frac{1}{2}V\{x^{**},p\}$, by construction, and by Lemma 1.
(ii) $x_i^{**}=CE\{(x_i^*,x_i^{**});(\frac{1}{2},\frac{1}{2})\}$ all i.

Hence if $\hat{x}_1^*=CE\{x^*,p\}$, and $\hat{x}_2^*=CE\{x^{**},p\}$,

$$x_0\{Ix,p\}I\{(\hat{x}_1^*,\hat{x}_2^*);(\tfrac{1}{2},\tfrac{1}{2})\}, \quad \text{by the inductive hypothesis}$$

$$V(CE\{x,p\})=\tfrac{1}{2}V(x_1^*)+\tfrac{1}{2}V(x_2^*)$$

$$=\tfrac{1}{2}V\{x',p\}+\tfrac{1}{2}V\{x,p\}=V\{x,p\}.$$

References

Allais, Maurice, 1952, Fondemonts d'une theorie positive des choix compartant un risque et critique des postulats et axiomes de l'ecole americaine, Paper presented at Colloque sur les Fondements et applications de la theorie du risque, Davis.

Allais, Maurice, 1953, Le comportement de l'homme rationnel devant le risque: Critique des postulats et axioms de l'ecole Americaine, Econometrica 62, 503–546.

Anderson, Jock, John Dillon and Brian Hardaker, 1977, Agricultural decision analysis (Iowa State University Press, Ames, IA).

Coombs, C.H., 1975, Portfolio theory and the measurement of risk, in: M. Kaplan and S. Schwartz, eds. (Academic Press, New York).

Dugundji, James, 1966, Topology (Allen and Bacon, Boston, MA).

Edwards, W., 1962, Subjective probabilities inferred from decisions, Psychological Review 69, 109–135.

Fellner, W., 1961, Distortion of subjective probabilities as a reaction to uncertainty, Quarterly Journal of Economics 75, 670–699.

Fishburn, Peter, 1964, Decision and value theory (Wiley, New York).

Fishburn, Peter, 1978, On Handa's 'New theory of Cardinal utility' and the maximization of expected return, Journal of Political Economy 86, 321–324.

Friedman, Milton and Leonard J. Savage, 1948, The utility analysis of choices involving risk, Journal of Political Economy 56, 279–304.

Hadar, Joseph and William Russell, 1969, Rules for ordering uncertain prospects, American Economic Review 59, 25–34.

Handa, Jagdish, 1977, Risk, probabilities and a new theory of cardinal utility, Journal of Political Economy 85, 97–122.

Hershey, J. and P. Schoemaker, 1980, Risk taking and problem context in the domain of losses: An expected utility analysis, Journal of Risk and Insurance 47.

Kahneman, Daniel and Amos Tversky, 1979, Prospect theory: An analysis of decision under risk, Econometrica 47, 263–293.

Karmarkar, U., 1978, Subjectively weighted utility: A descriptive extension of the expected utility model, Organisational Behavior and Human Performance 21, 61–72.

Karmarkar, U., 1979, Subjectively weighted utility and the Allais paradox, Organisational Behavior and Human Performance 24, 67–72.

Kunreuther, H., R. Ginsberg, L. Miller, P. Slovic, B. Botkan and N. Katz, 1978, Disaster insurance protection: Public policy lessons (Wiley, New York).

MacCrimmon, K., 1968, Descriptive and normative implications of the decision theory postulates, in: K. Borch and J. Mossin, eds., Risk and uncertainty (MacMillan, London).

Officer, R. and A. Halter, 1968, Utility analysis in a practical setting, American Journal of Agricultural Economics 50, 257–277.

Schoemaker, P. and H. Kunreuther, 1979, An experimental study of insurance decisions, Journal of Risk and Insurance 46, 603–618.

Slovic, P. and A. Tversky, 1974, Who accepts Savages axiom? Behavioral Science 19, 368–373.

Slovic, P., B. Fischhoff, S. Lichtenstein, B. Corrigan and B. Combs, 1977, Preference for insuring against probable small losses: Insurance implications, Journal of Risk and Insurance 44, 237–258.

Tesfatsion, L., 1976, Stochastic dominance and the maximization of expected utility, Review of Economic Studies 43, 301–316.

Tversky, Amos, 1969, Intransitivity of preferences, Psychological Review 76, 31–48.

Von Neumann, John and Oskar Morgenstern, 1944, Theory of games and economic behavior (Princeton University Press, Princeton, NJ).