Journal of Risk and Uncertainty, 1: 305–332 (1988)

Rank-Dependent, Subjective Expected-Utility Representations

R. DUNCAN LUCE
Harvard University

(**APPENDIX: Kenneth L. Manders,** *University of Pittsburgh*)

Key words: rank-dependent utility, uncertain payoffs, gambles, preference, dual-bilinear utility

Abstract

Gambles are recursively generated from pure payoffs, events, and other gambles, and a preference order over them is assumed. Weighted average utility representations are studied that are strictly increasing in each payoff and for which the weights depend both on the events underlying the gamble and the preference ranking over the several component payoffs. Basically two results are derived: a characterization of monotonicity in terms of the weights, and an axiomatization of the representation. The latter rests on two important conditions: a decomposition of gambles into binary ones and a necessary commutativity condition on events in a particular class of binary gambles. A number of unsolved problems are cited.

As a theory of choice between pairs of either risky or uncertain alternatives, subjective expected utility (SEU) is widely acknowledged to be normatively compelling, but not fully descriptive of behavior. Much that is wrong with SEU was anticipated early by Allais (1952/1979, 1953) and Ellsberg (1961), but relatively massive amounts of additional evidence were accumulated during the 1970s and early 1980s. Some of the data, especially those having to do with how the gambles are framed and the evidence for intransitivities (preference reversal phenomenon), went considerably beyond the earlier work. For an up-to-date statement of the issues from an economic perspective, see Machina (1987) and Weber and Camerer (1987), and for detailed list of the references concerning empirical tests, see Segal (1987a, p. 194).

One obvious possibility is to generalize the representation, keeping some of its attractive axiomatic features but omitting those that seem to be the source of major difficulty, as in Chew (1983), Fishburn (1983), Kahneman and Tversky (1979), Karmarkar (1978), Machina (1982), Quiggin (1982), Yaari (1987), Chew and Epstein (1987a), Gilboa (1987), Röell (1987), Segal (1987a, b, c), and Weber and Camerer (1987). The last of these references is a useful survey of all these developments. And Lola Lopes (personal communication) has noted Allais (1984, 1988a), which review his related approach to the problem.

From the point of view of the research to be reported here, a significant feature of most of this literature is that the domain of preferences is taken to be random variables on the line (Gilboa (1987) is an exception and Segal (1987a) is a partial exception). Although such a domain is very general and seems a relatively innocent way to model risky (although not uncertain) choices, its selection actually presupposes that there are no effects on preferences resulting from whether a gamble is framed as a succession of gambles or a single one with the same distribution. This assumption, which is inherent to the formulation as random variables, simply is contrary to a considerable body of data. It is important to understand why this is so. Since the objects of preference are (probability distributions of) random variables, it is not possible to distinguish between different ways of presenting the same random variable. For example, if x, y, z are sums of money and p, q, r are such that they and $p+q$ are in (0, 1), then the gamble presented as $(x, p; y, q; z, 1-p-q)$ is not distinguished as different from the mixture (or two-stage) gamble $(p+q)(x, p/(p+q); y, q/(p+q)) + (1-p-q)z$. Often this assumption is left implicit, although Segal (1987a, p. 181) states it explicitly. Obviously, if one is carrying out an experiment, these are distinctively different stimuli, and the evidence strongly suggests that subjects do not necessarily treat them as indifferent in preference. For this reason, it appears that theoretical developments based on taking random variables on the line as the domain are doomed to be descriptively inadequate.

An independent approach to the problem, assuming uncertainty rather than risk and avoiding this strong assumption of no framing effects, was given by Luce and Narens (1985). They provided a somewhat novel analysis of both the basic structure of the SEU axioms and of the implications of the empirical data for their accuracy.[1] Two aspects of their approach are unique.

The first is that the domain is definitely not random variables. Rather it is built up recursively from pure payoffs (possibly, money) and chance events combined to form chance mixtures that in turn are combined to form more complex chance mixtures. (Of course, random variables can also be constructed recursively from numerical payoffs and events with probabilities, but that is not usually done explicitly.) This recursive structure has a strong algebraic aspect that is exploited to deal with the class of framing effects mentioned above. Within this theory, if a payoff is contingent upon the joint occurrence of two independent events, we do not necessarily assume the decision-maker to be indifferent either to the order in which the events occur or to the single-stage gamble involving events whose probability of occurrence is equal to the probabilities of the corresponding payoffs in the two-stage gamble, i.e., the probabilities of their various joint occurrences. To some extent, Segal (1987a) has considered the same domain, and Gilboa (1987) considered the Savage-like nonrecursive state-space approach.

The second aspect of their approach that is different is that it does not provide a direct axiomatization in terms of properties of preference, but asks for the most general form of an interval-scale utility function—i.e., unique up to positive affine transformations—that is monotonically increasing in each payoff.

A third aspect, which is hardly a virtue, is that their results are limited to gambles with two payoffs that may themselves be gambles.

The present paper has two goals. The first is to attempt to generalize the ideas embodied in the Luce–Narens paper to gambles with more than two payoffs. The second is to develop a more conventional axiomatization of that representation in the sense that it involves only axioms about the primitives of the system and makes no explicit assumptions about the uniqueness of the utility representation. The resulting mathematical form of the representation bears a very close family resemblance to the rank-dependent models that have arisen in the random-variable case, but they are by no means identical, as will be shown.

1. Dual-bilinear utility

For binary gambles, Luce and Narens (1985) approached the problem as follows. Starting with a (sufficiently rich) set of pure payoffs (i.e., those not decomposable into gambles, e.g., sums of money) and a (sufficiently rich) set $\mathscr{E}$ of events from a sample space E, one generates recursively a gamble from each event A in $\mathscr{E}$ and each x and y, which are either gambles or pure payoffs. It is an event mixture of x and y in the sense that payoff x is received if event A occurs and payoff y is received if A fails to occur ($\bar{A}$). Such a gamble is denoted by $x\text{O}_A y$ The operator notation is intentional. Each occurrence of an event, including successive realizations of a single event as in $(x\text{O}_A y)\text{O}_A z$, is to be interpreted as independent of any other occurrence. To be completely precise, this really should be written $(x\text{O}_A y)\text{O}_{A'}$, where A and A' are two independent realizations of the same event, such as red coming up on two successive spins of the same roulette wheel. For notational simplicity, we drop the primes or subscripts and let the parentheses remind us that the realizations are independent.

Let $\mathscr{M}$ denote this family of gambles and assume $\succsim$ is a preference order over it ($\succ$ and $\sim$ are defined from $\succsim$ in the usual fashion). Furthermore, for each $A \in \mathscr{E}$, let $\mathscr{M}_A$ denote the subfamily of gambles obtained by starting with the pure payoffs and recursively using only the event A.

Let us assume the structure of gambles is sufficiently regular and rich that for each $A \in \mathscr{E}$, $<\mathscr{M}_A, \succsim_A>$, where $\succsim_A$ is the restriction of $\succsim$ to $\mathscr{M}_A$, meets the following conditions: for every $x, y, z, w \in \mathscr{M}_A$,

M1. (Weak order): $\succsim_A$ is transitive and connected (i.e., complete).
M2. (Strict Monotonicity): $x \succsim_A y$ iff $x\text{O}_A z \succsim_A y\text{O}_A z$ iff $w\text{O}_A x \succsim_A w\text{O}_A y$.
M3. (Weak Idempotence): $x\text{O}_A x \sim_A x$.
M4. (Solvable): there exist $u, v \in \mathscr{M}_A$ such that $x\text{O}_A u \sim_A y \sim_A v\text{O}_A x$.
M5. (Dedekind complete): any subset of $\mathscr{M}_A$ that is bounded from above relative to $\succsim_A$ has a least upper bound in $\mathscr{M}_A$.

The first three axioms are, of course, substantive, but they can be argued to hold for preference judgments (see below). The last two are structural and simply insure a sufficiently rich situation so as to get a representation onto the real numbers.

Solvability, as stated, implies that the structure is unbounded in preference. This is to some degree unrealistic and must be viewed as an idealization of the model.

The weak-order axiom, M1, can be questioned in both of its aspects, and it has been so questioned both normatively (Anand, 1987) and empirically. As was remarked in footnote 1, it is not apparent that there really exists strong empirical evidence against transitivity of choices (except possibly for nontransitivity of indifference of gambles that differ very little), although beyond a doubt combining choices with indifference judgments does sometimes lead to intransitivities. In empirical studies of simple gambles, at least, connectedness is not a concern, although some believe that it may be in complex real-world decisions.

The discussion of M2, strict monotonicity, requires a rather long aside because the issues of relating it to more familiar assumptions are not completely transparent. If one views the operation $\bigcirc_A$ as a function from $\mathcal{M}_A \times \mathcal{M}_A$ into $\mathcal{M}_A$, then for a mathematician it seems quite natural to speak of M2 as strict monotonicity (relative to $\succsim_A$), and this was the terminology adopted some time ago in the axiomatic measurement literature. However, it may generate confusion among economists who use monotonicity in other ways, such as for first-order stochastic dominance. From their perspective, M2 appears to be formally the same as the independence axiom, namely

M2′. If F, G, H are distributions and $p \in (0, 1)$, then $F \succsim G$ iff $pF + (1 - p)H \succsim pG + (1 - p)H$.

M2 is indeed formally the same, but only in a qualified sense since M2′ is used with random variables and so $pF + (1 - p)H$ is interpreted as a random variable, not as a distinctive entity called a probability mixture of F and H. Put another way, if payoff x can arise both in the gamble F and in the gamble H, that fact is irrelevant and all that counts is the total probability with which x arises. Such an identification is neither intended nor made in the statement M2.

To gain some idea of how significantly different this may be, consider another property whose violation is often called the Ellsberg paradox.

M6. Suppose $A, B, C \in \mathscr{E}$ are such that $A \cap C = \emptyset = B \cap C$ and $x \succ y$. Then $x \bigcirc_A y \succsim x \bigcirc_B y$ iff $x \bigcirc_{A \cup C} y \succsim_A x \bigcirc_{B \cup C} y$.

The comparable axiom in random variable notation is

M6′. Suppose G and H are random variables, $p, q, r, p + r, q + r \in (0, 1)$, and $G \succ H$. Then,
$pG + (1 - p)H \succsim qG + (1 - q)H$ iff $(p + r)G + (1 - p - r)H \succsim (q + r)G + (1 - q - r)H$.

As Luce and Narens (1985) demonstrated, M6 is a property independent of the remaining axioms, in particular, of M2. In contrast, as is well known, within the domain of random variables, M6′ follows from M2′. Such distinctions are, therefore, decidedly important in evaluating the adequacy of the current model to accommodate the empirical findings that have undercut SEU theory.

If one does not impose an axiom that reduces compound gambles, such as those that arise in M2, to one-stage, logically equivalent ones, then the little data that do exist plus most experimenters' intuitions lead one to accept M2 as empirically valid. It should, however, be carefully checked empirically.

Idempotence, M3, asserts the absence of one potential framing effect, one that is difficult to deny if the meaning of the gamble $x\bigcirc_A x$ is clear to the decision maker.

In addition to the assumptions M1–M5 about the behavior of the preference order, Luce and Narens (1985) made the following (automorphism) assumption that is of a quite different character.

A1. For each A, $<\mathscr{M}_A, \succsim_A>$ is a homogeneous structure in the sense that there exists an automorphism (isomorphism of the structure with itself) that takes each gamble into each other gamble.

This axiom entails both a certain richness and symmetry to the entire set of gambles. A more detailed summary of the use of automorphisms in the context of measurement can be found in Luce and Narens (1987).

From assumptions M1–M5 and A1, Luce and Narens (1985) were able to show that there exists a function U_A from $\mathscr{M}_A$ onto the positive reals, Re^+, and a function f_A from Re^+ onto Re^+ such that

1. (Order preserving): $x \succsim_A y$ iff $U_A(x) \geqslant U_A(y)$,
2. f_A is strictly increasing,
3. $f_A(r)/r$ is a strictly decreasing function of r, and
4. $U_A(x\bigcirc_A y) = U_A(y) f_A[U_A(x)/U_A(y)]$.

In order to arrive at a utility theory to cover all binary gambles, they made the following additional assumptions:

A2. By an appropriate normalization, it is possible to suppress the subscript A from U_A, i.e., for all $A, B \in \mathscr{E}$, $U_A = U_B = U$.

Below a qualitative property is provided that is equivalent to A2.

A3. The class of all utility functions forms an interval scale, i.e., all equally satisfactory utility functions are related to U by the positive linear transformations $U \rightarrow rU + s$, where r and s are real and $r > 0$.

From the representation arising from M1–M5 and A1, they were able to establish from A2 and A3 that the numerical operation given by f_A has a simple bilinear form—namely, there exist two functions, $S^+, S^-: \mathscr{E} \to (0, 1)$, such that, for all x, $y \in \mathscr{M}_A$ and $A \in \mathscr{E}$,

$$U(x \bigcirc_A y) = \begin{cases} U(x)S^+(A) + U(y)[1 - S^+(A)], & U(x) \geqslant U(y) \\ U(x)S^-(A) + U(y)[1 - S^-(A)], & U(x) \leqslant U(y), \end{cases} \tag{1}$$

Luce and Narens spoke of equation (1) as the *dual-bilinear* theory of utility. Note that, as with the work of Quiggin (1982), Gilboa (1987), and Yaari (1987), the key aspect of the representation is that it is rank-dependent—an SEU type of model with the weights assigned to events dependent upon their ranking according to the perceived value of the payoffs. Observe that equation (1) reduces to SEU when $S^+ \equiv S^- \equiv$ a probability measure. They explored carefully the conditions under which such a reduction occurs. The gist of the answer is that $S^+ \equiv S^-$, i.e., rank independence, follows from an assumption that certain rather simple pairs of gambles that are functionally the same are also judged to be indifferent in preference. *Functionally the same* means that the two gambles yield the same payoff under the same conditions except for the order in which successive events are realized. These equivalences are called *accounting equations*, and essentially they deny, contrary to data, the existence of certain types of the framing effects discussed by Kahneman and Tversky (1979) and Tversky and Kahneman (1986). An example of one that forces $S^+ \equiv S^-$ is bisymmetry:

$$(x \bigcirc_A y) \bigcirc_A (u \bigcirc_A v) \sim_A (x \bigcirc_A u) \bigcirc_A (y \bigcirc_A z).$$

If, in addition, the common weighting function is also to be a finitely additive probability, a more substantive assumption is needed such as the extended sure-thing property that was formulated as M6 above, and whose violation is, as noted earlier, commonly called the Ellsberg paradox.

Dual-bilinear utility can be written in another, equivalent way:

$$U(x \bigcirc_A y) = U(x)P(A) + U(y)[1 - P(A)] + W(A)|U(x) - U(y)|, \tag{2}$$

where $P(A) = \frac{1}{2}[S^+(A) + S^-(A)]$ and $W(A) = \frac{1}{2}[S^+(A) - S^-(A)]$. Since the standard deviation of U relative to $P(A)$ is

$$P(A)^{1/2}[1 - P(A)]^{1/2}|U(x) - U(y)|, \tag{3}$$

one can think of the model as saying that for binary gambles, utility of a gamble is a linear combination of the mean and standard deviation of the utilities of its components relative to a subjective probability distribution.

The issue to be considered in the remainder of the paper is how to generalize the dual-bilinear representation to gambles with more than two payoffs. It is by no means obvious how best to do this. One way to proceed is to try to reduce an n-payoff gamble to a sequence of concatenated two-payoff gambles, and so reduce them to something we can calculate from the dual-bilinear theory. In doing this, however, we must take care not to unwittingly assume accounting equations that are known to be empirically dubious. One possible way is shown below.

2. Invariance under interval scale transformations

Let n be an integer and let Π_n be the set of all ordered partitions of E into nontrivial events of $\mathscr{E}$. Suppose $\pi \in \Pi_n$, i.e., $\pi = (A_1, \ldots, A_n)$, where $A_i \in \mathscr{E}, A_i \cap A_j = \emptyset$ for $i \neq j$, $i, j = 1, \ldots, n$, and $\cup_i A_i = E$. Then a gamble over π is a generalized operation[2] $F_\pi(x_1, \ldots, x_n)$ that is interpreted as follows: $F_\pi(x_1, \ldots, x_n)$ denotes the one-stage gamble in which x_i is the payoff when A_i occurs. In particular, if $\pi = \{A, \bar{A}\}$, then $F_\pi(x,y) = x \bigcirc_A y$. Let $\mathscr{M}$ denote a set of such gambles including all of the pure payoffs and let $\succsim$ be the preference order over them. We make the following assumptions:

G1. $\succsim$ is a weak order on $\mathscr{M}$.

G2. The operation F_π is strictly monotonic increasing with respect to the preference order in each argument, i.e., for every $i = 1, \ldots, n$, and $x_1, \ldots, x_i, \ldots, x_n, y_i \in \mathscr{M}$.

$$y_i \succsim x_i \text{ iff } F_\pi(x_1, \cdots, y_i, \cdots, x_n) \succsim F_\pi(x_1, \cdots, x_i, \cdots, x_n).$$

G3. The operation F_π is idempotent in the sense that for each $x \in \mathscr{M}$, $F_\pi(x, x, \ldots, x) \sim x$.

G4. There is a utility function $U: \mathscr{M} \to Re$ that is order-preserving and that is unique to positive linear transformations.

Introduce the abbreviations

$$\vec{x}_n = (x_1, \cdots, x_n),\ u_i = U(x_i), \text{ and } \vec{u}_n = (u_1, \cdots, u_n).$$

Then we may define the real function with real arguments, G_π, by

$$G_\pi(\vec{u}_n) = UF_\pi(\vec{x}_n).$$

By monotonicity, G_π is well-defined. Since, by assumption, U is unique up to an interval scale, it follows that G must satisfy the functional equation

$$rG_\pi(\vec{u}_n) + s = G_\pi(r\vec{u}_n + s), \qquad (r, d \in Re, r > 0). \tag{4}$$

So the problem is to find solutions G_π to equation (4) that are idempotent and strictly increasing in each argument.

The assumption of monotonicity, G2, is sometimes, in this context, called the sure-thing principle. The reformulation of G2 in terms of the function G_π was refered to by Allais (1952/1979) as the axiom of absolute preference. Theorem I of Allais (1988b) explicitly formulates equation (4) and derives a form of solution (see the next section) that he had used implicitly earlier in Allais (1979, 1984).

3. Moment solutions to equation (4)

Such invariance as embodied in equation (4), but usually without the restriction $r > 0$, arises in many problems. For example, Chew, Epstein, and Segal (1987) investigated a special class of solutions, and Aczél (1966) reports a number of results. However, finding all monotonic increasing solutions to equation (4) apears to be difficult, in part at least, for the following reason. Aczél, Roberts, and Rosenbaum (1986) considered equation (4) with $r \neq 0$ rather than $r > 0$, and they pointed out that it trivially implies

$$G_\pi(\vec{u}_n) = M(\vec{u}_n) + S(\vec{u}_n)G_\pi[[\vec{u}_n - M(\vec{u}_n)]/S(\vec{u}_n)], \tag{5}$$

where M and S are the mean and standard deviation of the uniform distribution over $\vec{u}_n$. They did not mention, although it is equally true, that equation (5) is formally also a solution to equation (4) for many other choices of M and S. All that is required is that they transform as $M(r\vec{u}_n) = rM(\vec{u}_n) + s$ and $S(r\vec{u}_n) = rS(\vec{u}_n)$. Thus, they may be selected to be the mean and standard deviation of any distribution, or as any dimensional constants that transform appropriately, as often occurs in the formulation of physical laws.

Another family of solutions, investigated by Allais (1979, 1984), Munera (1985, 1986), and Munera and Neufville (1983), involves weighted sums of the k^{th} roots of the absolute values of the k^{th} moments about some mean. Solutions of this type can be thought of as plausible generalizations of the dual-bilinear model, at least when it is presented in its second form, equations (2) and (3).

However, there are two difficulties in using any of these solutions:

(i) They are for the general class of linear transformations $U \rightarrow rU + s$, $r \neq 0$, rather than the order-preserving ones, which necessitates $r > 0$, corresponding to interval scales; and

(ii) It is exceedingly difficult to see the impact of strict monotonicity on these solutions—moments and monotonicity simply are not very comfortable companions.

Once one takes (i) into account, we see that equation (4) is not a single functional equation, but a collection of them—one for each possible order of the n arguments. Thus, the solution space is even more complex than suggested by the solutions mentioned.

Allais (1984, p. 39) seems to say that what is alleged in (ii) is not true, but it is not clear where he works out explicitly the impact of monotonicity (in his terms, the axiom of absolute preference) on a general function of the moments. This impact appears to be very complex. There are, however, other formulations of the solutions to equation (4) that are somewhat more amenable to the imposition of monotonity (see the fifth unsolved problem discussed in section 9).

Since it is not obvious how to pursue the general class of moment solutions effectively, another approach follows.

4. Decomposition postulate

Despite the fact that monotonicity and interval scale transformations are sufficient to pin down the form in the binary case, they are not sufficient in the general case. We appear to need an additional assumption; however, considerable care must be exercised not to impose so much structure that matters reduce to SEU. An answer is suggested by theorem 3, p. 237 of Aczél's (1966) book.

Since we shall be considering each order of payoffs separately, there is no loss of generality in assuming the arguments have been permuted so that $x_1 \precsim x_2 \precsim \ldots \precsim x_n$.

Consider the following *decomposition postulate*:

G5. For $n > 2$, given $\pi \in \Pi_{n+1}$ and F_π on $n + 1$ arguments, there exist partitions $\sigma \in \Pi_n$ and $\gamma \in \Pi_2$ such that for all $x_1 \precsim \cdots \precsim x_n \precsim x_{n+1}$,

$$F_\pi(x_1, \cdots, x_n, x_{n+1}) \sim F_\eta[F_\gamma(x_1, \cdots, x_n), x_{n+1}]. \tag{6a}$$

Assuming a utility function exists, as above, (6a) immediately translates into the following: for all $u_1 \leqslant \ldots \leqslant u_n \leqslant u_{n+1}$,

$$G_\pi(u_1, \cdots, u_n, u_{n+1}) = G_\eta[G_\gamma(u_1, \cdots, u_n), u_{n+1}]. \tag{6b}$$

Axiom G5 simply says that any gamble on $n + 1$ payoffs is indifferent to a gamble having just two payoffs, the first of which is itself a gamble composed of the n payoffs other than the most preferred and the second of which is the most preferred one. There is no implication in this assumption that an accounting equation holds in terms of the objective probabilities of the payoffs occurring, although it clearly says just that in terms of some form of subjective weighting of the payoffs.

This property is similar to, but a little weaker than, the one Aczél invoked on p. 237.

Theorem 1. *Suppose the family of generalized numerical operations has a common utility function such that*

(i) *for each subfamily of binary operations generated by a single event, the utility function yields a dual bilinear representation of that subfamily;*

(ii) *each generalized operation is strictly increasing in each argument* (*see* G2);
(iii) *each generalized operation is idempotent* (*see* G3); *and*
(iv) *the family satisfies the decomposition postulate* (*see* G5).

Let ρ *denote the permutation of* $\{1, \ldots, n\}$ *such that*

$$u_{\rho(1)} \leqslant u_{\rho(2)} \leqslant \cdots \leqslant u_{\rho(n)}.$$

Then

(1) *There exist constants* $a_i(\pi, \rho)$ *such that*

$$G_\pi(\vec{u}_n) = \sum_{i=1}^{n} a_i(\pi, \rho) u_i, \tag{7a}$$

$$a_i(\pi, \rho) \geqslant 0 \text{ and} \sum_{i=1}^{n} a_i(\pi, \rho) = 1. \tag{7b}$$

(2) *Each generalized operation is invariant under positive affine transformations* (*equation* (4)).
(3) *Suppose* ρ *and* ρ' *are two permutations for which* u_i *and* u_j *are adjacent and the permutations are identical except that the order of* u_i *and* u_j *is reversed. Then*

$$a_k(\pi, \rho) = a_k(\pi, \rho'), \qquad k \neq i, j. \tag{8}$$

Proof (1) By hypothesis (i), equation (1) holds for $n = 2$, which is equation (7) with a change of notation for the weights. We prove the general case by induction. Note, first, that the decomposition postulate is compatible with the invariance hypothesis. Accordingly, by the induction hypothesis, we know that there are constants c, $0 < c < 1$, and $b_1, \ldots, b_n$, $b_i > 0$, $\Sigma_i b_i = 1$ such that

$$G_\sigma(u_1, \cdots, u_n) = \sum_{i=1}^{n} b_i u_i,$$

$$G_\eta(v, u_{n+1}) = cv + (1 - c)u_{n+1}, \qquad v \leqslant u_{n+1}.$$

By monotonicity and idempotence,

$$G_\sigma(u_1, \cdots, u_n) \leqslant G_\sigma(u_n, \cdots, u_n) = u_n \leqslant u_{n+1}.$$

Substituting into equation (6b), we see

$$G_\pi(u_1, \cdots, u_{n+1}) = \sum_{i=1}^{n+1} a_i u_i,$$

where $a_i = cb_i, i = 1, \ldots, n$, and $a_{n+1} = 1 - c$. Since we have assumed a permutation to put us in a standard situation, the constants a_i clearly are a function of both the partition, π, of the events underlying the payoffs as well as the ordering of the payoffs, ρ.

(2) Observe that the form given in equation (7a) satisfies the invariance property of equation (4).

(3) There remains the question of restricting the solutions to those that are monotonic in each variable. Monotonicity is exhibited by a weighted mean so long as the change in the variable does not alter the order among the payoffs. However, it need not be exhibited when the order is changed, in which case a discontinuous change occurs in the weights assigned to some events. We show that, given equation (7), such monotonicity is equivalent to equation (8).

With no loss of generality, suppose $u_i < u_j$ and they are adjacent. Consider what happens as u_i is moved up and beyond u_j, say $u_i' < u_j < u_i''$, but with the upper value not crossing the next larger value. Thus, by theorem 1 and assuming monotonicity,

$$\sum_{k \neq i,j} a_k(\rho)u_k + a_i(\rho)u_i' + a_j(\rho)u_j \leqslant \sum_{k \neq i,j} a_k(\rho')u_k + a_i(\rho')u_i'' + a_j(\rho')u_j. \quad (9)$$

Similarly, select $u_j' < u_i < u_j''$, and monotonicity yields

$$\sum_{k \neq i,j} a_k(\rho')u_k + a_i(\rho')u_i + a_j(\rho')u_j' \leqslant \sum_{k \neq i,j} a_k(\rho)u_k + a_i(\rho)u_i + a_j(\rho)u_j''. \quad (10)$$

In equation (9) let u_i' and u_i'' approach u_j and in equation (10) let u_j' and u_j'' approach u_i. This yields

$$\sum_{k \neq i,j} a_k(\rho)u_k + [a_i(\rho) + a_j(\rho)]u_j \leqslant \sum_{k \neq i,j} a_k(\rho')u_k + [a_i(\rho') + a_j(\rho')]u_j, \quad (11)$$

$$\sum_{k \neq i,j} a_k(\rho')u_k + [a_i(\rho') + a_j(\rho')]u_i \leqslant \sum_{k \neq i,j} a_k(\rho)u_k + [a_i(\rho) + a_j(\rho)]u_i. \quad (12)$$

Adding equations (11) and (12) and rearranging,

$$[a_i(\rho) + a_j(\rho)](u_j - u_i) \leqslant [a_i(\rho') + a_j(\rho')](u_j - u_i).$$

Since $u_j > u_i$, we see $a_i(\rho) + a_j(\rho) \leqslant a_i(\rho') + a_j(\rho')$. The opposite inequality can be achieved by a parallel argument beginning with ρ' and $u_j < u_i$, and so

$$a_i(\rho) + a_j(\rho) = a_i(\rho') + a_j(\rho'). \quad (13)$$

Substituting equation (13) into equations (11) and (12), we see that

$$\sum_{k \neq i,j} a_k(\rho)u_k = \sum_{k \neq i,j} a_k(\rho')u_k. \tag{14}$$

Since the u_k are free to vary, subject to the inequalities imposed by ρ and ρ', equation (14) implies $a_k(\rho) = a_k(\rho')$, $k \neq i, j$.

Strict increasing monotonicity follows immediately from this condition. Q.E.D.

This representation generalizes the dual-bilinear model by maintaining the fact that the dual-bilinear model is an expectation for each possible order of the payoffs. It does not generalize the fact that the binary model can also be viewed as a weighted sum of an expectation and standard deviation. This representation was postulated by Sen (1973) and also by Chew and Epstein (1987b), the latter in their study of Gini indices.

Although monotonicity reduces the number of independent parameters, i.e., the possible $\{a_i(\cdot)\}$, for fixed n down from the potential $(n-1)n!$, there are still very many. K.L. Manders (see Appendix) has shown that there are $2^n - 2$. If, however, the decomposition property, equation (6), is understood in detail, then of course the only parameters that need be estimated are the binary ones.

For fixed n, a further reduction is achieved by assuming that the weight associated to payoff i when the permutation ρ holds depends only on the rank of i, not on how the other alternatives are ranked. Stated formally, if $\rho(i) = r$, then $a_i(\rho) = a_i(r)$.

Corollary. *Suppose the conditions of Theorem* 1 *are met and suppose that for* $\rho(i) = r$, $a_i(\rho) = a_i(r)$. *Then for all* i, j,

$$a_i(r) = a_j(r). \tag{15}$$

Proof. Let the weight associated with u_i when it is in rank position r be denoted by $a_i(r)$. By the theorem,

$$a_i(r) + a_j(r+1) = a_i(r+1) + a_j(r).$$

By induction, we show

$$a_i(r) + a_j(s) = a_i(s) + a_j(r). \tag{16}$$

With no loss of generality, suppose $r < s$. For $s = r + 1$ it has been proved. By the induction hypothesis on s,

$$a_i(r) + a_j(s-1) = a_i(s-1) + a_j(r),$$

and

$$a_i(s-1) + a_j(s) = a_i(s) + a_j(s-1).$$

Adding yields equation (16). Summing over $s \neq r$ in equation (16),

$$(n - 1)a_i(r) + \sum_{s \neq r} a_j(s) = \sum_{s \neq r} a_i(s) + (n - 1)a_j(r).$$

Using $\Sigma_{s=1}^{n} a_k(s) = 1$,

$$(n - 1)a_i(r) + 1 - a_j(r) = 1 - a_i(r) + (n - 1)a_j(r),$$

whence $a_i(r) = a_j(r)$. Q.E.D.

Observe that since there are n ranks and the weights sum to 1, this reduces the situation to $n - 1$ parameters. Thus, this specialization of the model has the same number of parameters as SEU, but except for the case of $n = 2$ it is totally different. It describes a person who totally ignores the partition and weights each payoff only according to its preference order among all payoffs, whereas SEU says the permutation is irrelevant, and only the partition of events matters.

5. An example of the constraints imposed by equations (8) and (15)

The simplest case for which equation (8) has any bite is, of course, for $n = 3$. Denote the utilities of the payoffs by u, v, and w. And suppose, for example, the order ρ is $u > v > w$. If the order is changed to $v > u > w$, i.e., u and v are interchanged, then the theorem tells us that the weight on the unchanged element, w, must be the same. Thus, we obtain the matrix of weights shown in table 1.

Thus, we see that from there being potentially 12 free parameters, monotonicity reduces the number of independent ones to six. That is still a lot of parameters, and as was noted the number rises rapidly with n, namely, as $2^n - 2$. The conditions of the corollary, i.e., that the weight attached to a payoff depends only upon its rank, yield the added constraints: $a = d = f$ and $b = c = e$. Thus, as was remarked earlier, there are only two parameters.

Table 1. Matrix of weights for the case $n = 3$

	u	v	w
$u > v > w$	a		e
$u > w > v$	a	c	
$v > u > w$		d	e
$v > w > u$	b	d	
$w > u > v$		c	f
$w > v > u$	b		f

(The blanks are filled in by the fact that the weights in each row sum to 1.)

6. Relations to rank-dependent theories of random variables

A natural question to raise is: How does the representation of theorem 1 relate to the weighting functions that have arisen in the rank-dependent theories arising on the assumption that the domain of gambles can be described as random variables? Recall that they are of the following form: there is a probability distribution $\vec{p}\ (\pi)$ over the partition π and a strictly increasing function φ such that

$$a_i(\pi, \rho) = \varphi\left[\sum_{\rho(j)\leqslant i} p_j(\pi)\right] - \varphi\left[\sum_{\rho(j)<i} p_j(\pi)\right]. \tag{17}$$

It is easy to verify that these weights exhibit the property of equation (8) since permutations that leave i fixed and do not permute something on one side of i to the other side leave the sums in equation (17) invariant.

What is not so obvious is whether or not for any rank-dependent SEU theory of the type described in theorem 1, there exists an increasing function φ and a family of probability distributions over the several partions such that equation (17) holds. Clearly, this would be false if we could derive from equation (17) some properties of the weights a_i not derivable from equation (8). This we do by looking at the $n = 3$ case in detail.

Suppose there were a strictly increasing function φ from the unit interval onto the unit interval for which equation (17) holds for any partition into three events, and so for any system of weights satisfying equation (8). Then referring to the general table of weights for the $n = 3$ case and denoting the unkown probabilities by p_u, p_v, and p_w, it is not difficult to check that equation (17) reduces to

$$a = 1 - \varphi(p_v + p_w), \qquad b = \varphi(p_u), \qquad c = \varphi(p_v),$$
$$d = 1 - \varphi(p_u + p_w), \qquad e = \varphi(p_w), \qquad f = 1 - \varphi(p_u + p_v).$$

From these, it follows immediately that φ^{-1} must meet the following three constraints:

$$\varphi^{-1}(1 - a) = \varphi^{-1}(c) + \varphi^{-1}(e),$$
$$\varphi^{-1}(1 - d) = \varphi^{-1}(b) + \varphi^{-1}(e),$$
$$\varphi^{-1}(1 - f) = \varphi^{-1}(b) + \varphi^{-1}(c).$$

The six parameters are independent, but subject to the constraints that they are positive and

$$1 - a > c \text{ and } e,\ 1 - d > b \text{ and } e, \text{ and } 1 - f > b \text{ and } c.$$

Since φ is strictly increasing and onto, it is continuous. So if we let c and e approach $1 - a$, in the limit we must have for any choice of $1 - a$, $\varphi^{-1}(1 - a) =$

$2\varphi^{-1}(1 - a)$, which is impossible. Thus, in general, the rank-dependent weights cannot be assumed necessarily to arise from a model of the form of equation (17).

7. Qualitative axiomatization of rank-dependent utility

The line of argument taken so far, although axiomatic and qualitative in character, departs rather shaply from the traditional approach to utility representations. The usual tack—Allais (1979) excepted—is to formulate the axiomatization directly in terms of the primitives, with no reference to the nature of the automorphism group, as has been done here by invoking interval scalability, which implies homogeneity.

Most of the existing attempts at axiomatizing rank-dependent utility have begun with a risk structure in which gambles can be thought of as random variables, in the sense that a probability measure over events is assumed. The major exceptions to that assertion are Segal (1987a) and Gilboa (1987) who, however, only arrived at the representation subject to the constraint formulated in equation (17). Thus, it remains an open issue to formulate an axiomatic theory leading to the representation embodied in equations 7 and 8 (theorem 1).One such axiomatization is outlined here that, to some extent, builds upon an existing result. Because it first axiomatizes the case of just two-payoff gambles generated from a single event and then extends that through the decomposition postulate to general ones, it may be judged a bit cumbersome, It is to be hoped that a more elegant theory can be devised.

The strategy to be followed is, first, to consider the structure of gambles that can be formed from repeated, independent realizations of binary events, the qualitative structure of which was described earlier. Accepting the assumptions M1-M5, Luce (1986) provided axioms necessary and sufficient for the existence of a dual-bilinear representation, equation (1). These axioms, which are somewhat complex to state, are postponed until the rest of the development is outlined.

The resulting interval-scale utility function U must be subscripted to show its dependence on the event A. So the second step is to add an axiom designed to insure that U_A does not really depend on A, i.e., there is a single utility function U over all of $\mathcal{M}$ that both preserves the order $\succsim$ and for which equation (1) holds. Part of the proof of theorem 2 shows that a necessary and sufficient condition for there to be a single utility function is:

$$(x \bigcirc_A y) \bigcirc_B y \sim (x \bigcirc_B y) \bigcirc_A y, \qquad (x, y, z \in \mathcal{M},\ A, B \in \mathcal{E}). \tag{18}$$

This is one of the simplest examples of an accounting equation. It simply postulates that changing the order of occurrence of the events in the gamble $(x \bigcirc_A y) \bigcirc_B y$ does not affect preference. Observe that under both orders, x is the final payoff if both A and B occur; otherwise, y is the payoff. The only difference is whether A

occurs before or after B. Moreover, the basic structure of this two-stage gamble is rather easy to perceive because there are only two payoffs and only two events.

As is easily shown, equation (18) is a necessary property of the dual-bilinear model and, unlike any of the more complex accounting equations, it does not destroy the property of being rank dependent.

The final step in the development is simply to invoke the decomposition postulate, equation (6), and monotonicity of gambles in each argument.

Now the axioms on $\mathscr{M}_A$ are described that lead to the dual-bilinear representation. For that purpose, the subscript A is suppressed. The general idea of the axiomatization is to extend the operation $\bigcirc$ in two ways. One corresponds to extending the upper half of equation (1) so that it holds for all x and y; that extended operation will be denoted by $*$. The other corresponds to extending the lower half of equation (1) to hold for all x and y; that operation will be denoted by $*'$. Thus, $*$ agrees with $\bigcirc$ for $x \succ y$ and $*'$, with $\bigcirc$ for $x \prec y$.

The first postulate we need is designed to insure that such an extension is possible and unique. Once done, the remaining postulates are stated in terms of $*$ and $*'$, and they are necessary and sufficient to have a representation as weighted averages.

Define $*$ as

$$x * y = \begin{cases} x \bigcirc y & \text{if } x \succ y, \\ x & \text{if } x \sim y, \\ w & \text{if } x \prec y, \end{cases} \tag{19}$$

where w is an element that has the property that there exist u, v such that

$$u \succsim x, \; v \succsim y, \; u \bigcirc x \succsim v \bigcirc y, \; u \bigcirc v \succsim w, \tag{20}$$

and

$$(u \bigcirc x) \bigcirc (v \bigcirc y) \sim (u \bigcirc v) \bigcirc w. \tag{21}$$

One can prove from axioms M1–M4 that such an element exists (lemma 5.1 of Luce, 1986). The assumption that must be made is that for the w mentioned above and for any u and v meeting the four conditions of equation (20), then they also satisfy equation (21).

The operation $*'$ is defined in an analogous way to equation (19) and the analogue of the property leading to equation (21) must be assumed.

A structure exhibiting both of these properties is called *reflectable*. Note that any structure having a dual-bilinear representation is reflectable.

One shows that reflectableness together with axioms M1–M4 is enough to insure that $*$ and $*'$ are well-defined, idempotent, intern (i.e., for $x \succ y$, then $x \succsim x \bigcirc_A y \succsim y$), monotonic, and solvable (lemma 5.2 of Luce, 1986).

The remaining assumption are that $*$ and $*'$ are both *right autodistributive*, i.e., for $x, y, z \in \mathscr{M}_A$,

$$(x * y) * z \sim (x * z) * (y * z) \text{ and } (x *' y) *' z \sim (x *' z) *' (y *' z), \tag{22}$$

and that together they satisfy *generalized bisymmetry*, i.e., for every x, y, u, $v \in \mathscr{M}_A$,

$$(u * v) *' (x * y) \sim (u *' x) * (v *' y). \tag{23}$$

Theorem 2. *Let $<E,\mathscr{E}>$ be an event space, X a set of pure payoffs, $\mathscr{M}$ the set of gambles generated recursively from X and finite partitions of E into events of $\mathscr{E}$, and $\succsim$ a binary relation on $\mathscr{M}$. For each $A \in \mathscr{E}$, let $\mathscr{M}_A$ denote the set of gambles generated recursively from X and A. Suppose the following assumptions are met:*

(i) *For each $A \in \mathscr{E}$, $<\mathscr{M}_A, \succsim>$ satisfies axioms* M1–M5, *is reflectable,* (*see equations* (*20*) *and* (*21*) *and their analogues*) *and the defined operations $*$ and $*'$* (*equation* (*19*) *and its analogue*) *are both right autodistributive* (*equation* (*22*)), *and together satisfy generalized bisymmetry*) (*equation* (*23*));

(ii) *For every x, $y \in \mathscr{M}$ and A, $B \in \mathscr{E}$, equation* (*18*) *holds, i.e.,*

$$(x \bigcirc_A y) \bigcirc_B y \sim (x \bigcirc_B y) \bigcirc_A y; \text{ and}$$

(iii) *For each partition π of E into $n + 1$ events, there exists a partition σ of E into n events and an event A such that for each $x_i \in \mathscr{M}$, $i = 1, \ldots, n + 1$, with $x_1 \precsim x_2 \precsim \ldots \precsim x_{n+1}$, equation* (6a) *holds, i.e.,*

$$F_\pi(x_1, \cdots, x_{n+1}) \sim F_\sigma(x_1, \cdots, x_n) \bigcirc_A x_{n+1}. \tag{24}$$

Then there exists an interval scale representation of $<\mathscr{M}, \succsim>$ that has the form given in theorem 1.

Proof. Suppose $\mathscr{X} = <\mathscr{M}_A, \succsim, \bigcirc_A>$ is a concatenation structure that satisfies properties M1–M5. Luce (1986, theorem 5.1) has shown that $\mathscr{X}$ has a representation $<U_A, S^+, S^->$ of the dual-bilinear form given in equation (1) iff $\mathscr{X}$ is reflectable and the defined operations $*$ and $*'$ are right autodistributive and satisfy generalized bisymmetry.

Select any $x_1 \succ x_0$ and for each event A choose the representation U_A such that $U_A(x_1) = 1$ and $U_A(x_0) = 0$. This is possible because the representations of $<\mathscr{M}_A, \succsim_A, \bigcirc_A>$ form an interval scale. Furthermore, this representation is onto the real numbers because, by solvability, the structure is unbounded; it is order dense since, by idempotence and monotonicity, it is intern; and it is Dedekind complete by assumption. Thus, for each pair of events A and B there is a function G from the reals onto the reals such that $U_B = G(U_A)$. Moreover, since the U's are order preserving, G is strictly increasing. thus, G is continuous. The problem is simply to insure that it is the identity function. We now show that equation (18) is necessary and sufficient for that to be true.

Since the representations are onto the reals, applying $U_B = G(U_A)$ to the left side of equation (18) yields

$$\begin{aligned} U_B(x \mathrm{O}_A y)b + U_B(y)(1-b) &= G[U_A(x\mathrm{O}_A y)]b + G[U_A(y)](1-b) \\ &= G[U_A(x)a + U_A(y)(1-a)]b + G[U_A(y)](1-b), \end{aligned}$$

where a and b are the constants in equation (1) corresponding to events A and B, respectively, for whichever order is determined by x and y. Note that $x\mathrm{O}_A y$ lies between them, and so only the one constant will arise for each event.

Carrying out the same calculation for the right side yields

$$\begin{aligned} U_B[(x\mathrm{O}_B y)\mathrm{O}_A y] &= G[U_A[(x\mathrm{O}_B y)\mathrm{O}_A y]] \\ &= G[U_A(x\mathrm{O}_B y)a + U_A(y)(1-a)] \\ &= G[U_A U_B^{-1} U_B(x\mathrm{O}_B y)a + U_A(y)(1-a)] \\ &= G[G^{-1}[U_B(x)b + U_B(y)(1-b)]a + U_A(y)(1-a)] \\ &= G[G^{-1}[G[U_A(x)]b + G[U_A(y)](1-b)]a + U_A(y)(1-a)]. \end{aligned}$$

Setting $X = U_A(x)$ and $Y = U_A(y)$ and equating these two expressions, we see that G must satisfy the following functional equation:

$$\begin{aligned} G[Xa + Y(1-a)]b + G(Y)(1-b) = G[G^{-1}[G(X)b + G(Y)(1-b)]a \\ + \mathrm{Y}(1-a)]. \end{aligned} \tag{25}$$

Note that because the representations were selected to agree at 0 and 1, G must also satisfy $G(0) = 0$ and $G(1) = 1$. Setting $Y = 0$ in equation (25),

$$G(Xa)b = G[G^{-1}[G(X)b]a].$$

Letting $H(X) = G^{-1}[G(X)b]$, we see that $H(Xa) = H(X)a$. From this it follows readily that for integers m and n, $H(Xa^{m/n}) = H(X)a^{m/n}$. By the continuity of G, which arises from its being onto the reals and strictly increasing, H is also continuous. Thus, we may conclude for any real $\lambda > 0$, $H(XA^{\lambda}) = H(X)a^{\lambda}$. Setting $Z = a^{\lambda}$ and letting $X = 1$, we see that for $0 < Z < 1$, $H(Z) = cZ$, where $c = H(1) = G^{-1}(b)$. Substituting into the definition of H, $G(X)G(c) = G(X)b = GH(X) = G(cX)$. In exactly the same manner, $G(X)G(c^{\lambda})$ and so for $0 > X > 1$, $G(X) = X^{\beta}$.

Our task is reduced now to showing that $\beta = 1$. To that end, substitue G into equation (25) with $0 < X < 1$, $0 < Y < 1$, and introduce the variable $Z = X/Y > 0$:

$$(aZ + 1 - a)^{\beta}b + (1 - b) = [(Z^{\beta}b + 1 - b)^{1/\beta}a + (1 - a)]^{\beta}.$$

If this holds for all $Z > 0$, then it also holds for the derivative as $Z \to 0$. The left derivative approaches $\beta(1 - a)^{\beta-1}ab$ which is finite and >0 since a, b are in the open unit interval. The derivative of the right side involves a factor $Z^{\beta-1}$ multiplying positive terms, and so it approaches 0 or ∞ except when $\beta = 1$. Thus, $\beta = 1$, proving that the utility functions are identical over the interval x_0 to x_1. Since the argument can be made for any finite interval, we conclude that the utility functions must be identical.

The converse is trivial.

The conclusion is obvious from these facts and the decomposition assumption equation (24).

8. Finitely additive weighting functions for binary gambles

Much of utility theory involves expectations relative to a finitely additive probability measure. Thus, it is natural to ask whether within the context of rank-dependent utility theories there are reasonable conditions that lead one to expect the weighting functions to be finitely additive. Consider the simple binary gambles, where the representation becomes the dual-bilinear representation of equation (1). That equation is a special case of equation (7a) but with a different notation for the weights, which it is helpful to maintain since the binary case exhibits some properties that do not really generalize. For example, these weights depend upon just an event since that is equivalent to a partition; for $n > 2$, the dependence is on the partition, not just the event. So, for the binary but not the general case, it makes sense to ask about plausible conditions to force finite additivity, i.e. the condition:

$$\text{If } A, B \in \mathscr{E} \text{ and } A \cap B = \emptyset, \text{ then}$$
$$S^i(A \cup B) = S^i(A) + S^i(B), \quad (i = +, -). \tag{26}$$

Assuming that the rank-dependent model holds for binary gambles, the following two assumptions are sufficient to force equation (26).

FA1. For all $x, y \in \mathscr{M}$ and $A \in \mathscr{E}$,

$$x \bigcirc_A y \sim y \bigcirc_{\bar{A}} x. \tag{27}$$

As Luce and Narens (1985) noted, if the dual-bilinear model holds equation (27) is equivalent to:

$$S^+(A) + S^-(\bar{A}) = 1. \tag{28}$$

The second assumption is an appreciable strengthening of the decomposition postulate. The strengthening occurs in two respects. First, let us suppose that any gamble $F_\pi(x_1, \ldots, x_{n+1})$ on $n + 1$ events is equivalent in preference to a binary gamble having as one of its payoffs any x_i and the other a gamble involving the remaining n payoffs. Second, suppose further that the event used in the binary gamble is the complement of the event yielding x_i in the original gamble. This is plausible in that x_i is clearly isolated in the binary gamble and its relation to the events of the original gamble is transparent.

FA2. For any partition $\pi = \{A(1), \ldots, A(n+1)\}$ there exists for each $i = 1, \ldots, n + 1$ a partition $\sigma(i)$ such that

$$F_\pi(x_1, \cdots, x_i, \cdots, x_{n+1}) \sim F_{\sigma(i)}(x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_{n+1}) \bigcirc_{\bar{A}(i)} x_i. \tag{29}$$

In what follows, we really only need equation (29) for the case $n = 2$, in which case we may write $F_{\sigma(i)}(x, y) = x \bigcirc_{B(i)} Y$, and so it amounts to

$$\begin{aligned}(x_1 \bigcirc_{B(3)} x_2) \bigcirc_{\bar{A}(3)} x_3 &\sim (x_1 \bigcirc_{B(2)} x_3) \bigcirc_{\bar{A}(2)} x_2 \\ &\sim (x_2 \bigcirc_{B(1)} x_3) \bigcirc_{\bar{A}(1)} x_1.\end{aligned} \tag{30}$$

With no loss of generality, assume $x_1 \succ x_2 \succ x_3$. Note that the preference order between $x_1 \bigcirc_{B(2)} x_3$ and x_2 is not determined and, by axiom M4 together with the fact that $\mathscr{M}_A$ is mapped under U onto the real numbers, either order can actually occur. Applying the dual-bilinear model to the first and third equivalences and by varying x_1, we see that necessarily

$$S^+(B(3)) = S^+(A(1))/S^+(\bar{A}(3)). \tag{31}$$

Choosing x_1 so that $x_1 \bigcirc_{B(2)} x_3 \succ x_2$, applying the model to the first and second equivalences of equation (30), using equation (28), and varying x_2 yields

$$S^+(B(3)) = [S^+(\bar{A}(3)) + S^+(\bar{A}(2)) - 1]/S^+(\bar{A}(3)) \tag{32}$$

Equating equations (31) and (32), using equation (28) again, and noting that $\bar{A}(1) = A(2) \cup A(3)$, we see that

$$S^-(A(2) \cup A(3)) = S^-(A(2)) + S^-(A(3)),$$

which, since $A(2)$ and $A(3)$ may be selected to be any two disjoint events, is equation (26) with $i = -$.

The other case is similar: select x_3 sufficiently small so that $x_1 \bigcirc_{B(2)} x_3 < x_2$, vary x_2, and conclude

$$S^+(A(1) \cup A(2)) = S^+(A(1)) + S^+(A(2)),$$

i.e., finite additivity.

Equation (26) does not appear to have a natural generalization for gambles with more than two payoffs. The reason can be seen as follows. Suppose $x_1 \precsim x_2 \precsim \ldots \precsim x_{n+1}$ and suppose in equation (24) $A = \bar{A}(n + 1)$. Then, of course, the weight for x_{n+1} in the $(n + 1)$-payoff gamble is $1 - S^-(\bar{A}(n + 1)) = S^+(A(n + 1))$. By induction, the weight for x_n in the $(n + 1)$-payoff gamble is

$$S^-(\bar{A}(n + 1))[1 - S^-(\bar{A}(n)] = S^-(\bar{A}(n + 1))S^+(A(n)),$$

which depends upon more than just $A(n)$. Therefore, trying to state anything about how the weight of the union of events depends upon the two components is much complicated by its dependence upon the rest of the partition. There is no obviously easy way to formulate a result.

9. Discussion

Five questions arise about the model under discussion.

First, can a simpler, more elegant axiomatization of the order over gambles be arrived at that leads to the general rank-dependent representation? Certainly the one given her (theorem 2) requires a fairly complex pattern of assumptions, one group for binary gambles and quite a different one for more complex ones. Further, those for the binary case entail somewhat complex defined quantities. It would be desirable to find a more pleasing system.

Second, what conditions on the systems of weights are necessary and sufficient for them to be recovered from the rank-dependent model of equation (17) that has arisen in theories based on a domain of random variables? The answer to this question is not clear.

Third, there is a possible further constraint on weights that has not been investigated. If we think of the weights as analogous to subjective probabilities, but affected by the ordering of the payoffs, it might be plausible that the weights order the events independent of the ordering of the payoffs or, in the language introduced in this literature, that they be *co-monotonic* in the following sense: For any partition π, any rankings of payoffs ρ and σ, and any two events i and j of the partition,

$$a_i(\pi, \rho) \geqslant a_j(\pi, \rho) \text{ iff } a_i(\pi, \sigma) \geqslant a_j(\pi, \sigma). \tag{33}$$

What qualitative condition is equivalent to this apparently desirable property is unknown to the author. Note that it simply places inequalities on the weights but does not reduce their number. For example, in the case of $n=3$ shown in table 1, the following is an example of six independent parameters that meet the inequalities of equation (33): a = .48, b = .46, c = .40, d = .42, e = .06, f = .12.

Fourth, just how compatible are the decomposition and monotonicity properties? The former says that all weights arising in $n > 2$ cases are explicit functions of certain of the binary weights. The latter (via theorem 1) says that the weights in the $n > 2$ cases meet certain explicit constraints. It is by no means clear why the monotonicity constraints should hold given their generation by the decomposition postulate.

The last open problem is to gain a deeper understanding of all *monotonic* solutions to the invariance equation, equation (4). The problem seems to be complex. For example, consider the case $n = 3$, in which case we may write equation (4) as

$$G(ru + s, rv + s, rw + s) = rG(u, v, w) + s.$$

Setting $r = 1$ and $s = -w$, we see

$$G(u, v, w) = w + G(u - w, v - w, 0).$$

For $v \neq w$, set $r = 1/(v - w)$, $s = 0$, and $f(X) = G(X, 1, 0)$; then

$$G(u, v, w) = w + (v - w)f[(u - w)/(v - w)]$$

is a general solution to equation (4). Assuming the three partial derivatives of G exist, it is routine to verify that monotonicity is equivalent to: for real X,

$$f'(X) > 0,$$
$$f(X) - Xf'(X) > 0,$$
$$1 - f(X) + (X - 1)f'(X) > 0.$$

It is unclear how to characterize all solutions to this system of differential inequalities. An example of one that is different from the rank-dependent SEU model is:

$$f(X) = a + bX - c/(1 + e^{|X|}) \qquad (a > c/2 > 0, b > 0, 1 - a - b > 0).$$

This system of differential inequalities can be generalized in like manner to any $n > 3$.

Acknowledgments

The work reported here was supported, in part, by National Science Foundation grant IRI-8602765 to Harvard University. It was completed while the author was a Fellow at the Center for Advanced Study in the Behavioral Sciences, with financial support provided by National Science Foundation grant BNS-8700864 and by the Alfred P. Sloan Foundation. A preliminary version was presented at the annual meeting of the Society for Mathematical Psychology, Berkeley, CA, August, 1987, and a more complete one at the Joint National Meeting of the Institue of Management Sciences and the Operations Research Society of America, Washington, DC, April, 1988.

I have benefitted greatly from conversations with Mark Machina, who introduced me to the contemporary economic literature on utility with rank-dependent weights. I also thank Jean-Claude Falmagne for helpful comments on an earlier draft, A.A.J. Marley for several references, and an anonymous referee who made a number of valuable suggestions about the presentation.

Notes

1. As with all of the proposed generalizations, save for Fishburn (1982) and Loomes and Sugden (1982, 1983), Narens and I ignored the data that is interpreted as evidence against transitivity, mainly, the preference reversal phenomenon. (The original reference is Lichtenstein and Slovic, 1971; a major economic reference is Grether and Plott, 1979; and more recent references can be found in Bostic, Herrnstein, and Luce, 1988, and Tversky, Sattath, and Slovic, 1988.) One reason is that this phenomenon appears to arise largely because two quite different procedures are intermixed, namely, choices between gambles and the assignment of money values (i.e., certainty equivalents) to gambles. Evidence of Tversky et al. (1988) and of Bostic et al. (1988) (as well as unpublished research of Pearson, 1986), strongly suggest that this is in fact the primary source of the intransitivities. Since the theory presented here speaks only of choices, it is not clear that accommodation for intransitivity is really needed.

2. Aczél and Saaty (1983) have studied a very special operation that can be written as $g(x_1) \bigcirc g(x_2) \bigcirc \ldots \bigcirc g(x_n)$, where $\bigcirc$ denotes an associative operation. This is much too strong an assumption for the purposes of generalizing SEU.

References

Aczél, J. *Lectures on Functional Equations and Their Applications*. New York: Academic Press, 1966.

Aczél, J., Roberts, F.S. & Rosenbaum, Z. On Scientific Law without Dimensional Constants. *Journal of Mathematical Analysis and Applications*, (Vol. 119, 1986), pp 389-416.

Aczél, J. & Saaty, T.L. Procedures for Synthesizing Ratio Judgements. *Journal of Mathematical Psychology*, (Vol. 27, 1983), pp 93-102.

Allais, M. Foundations of a Positive Theory of Choice Involving Risk, and a Criticism of the Postulates and Axioms of the American School. In: M. Allais and O. Hagen eds. *Expected Utility Hypothesis and the Allais' Paradox*. Dordrecht: D. Reidel Publishing Co., 1952/1979.

Allais, M. Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine. *Econometrica* (Vol.21, 1953), pp 503-546.

Allais, M. The So-Called Allais' Paradox and Rational Decisions under Uncertainty. In: M. Allais and O. Hagen, eds, *Expected Utility Hypothesis and the Allais' Paradox*. Dordrecht: D. Reidel Publishing Co., 1979.

Allais, M. the Foundations of the Theory of Utility and Risk. In: O. Hagen and F. Wenstop, eds., *Progress in Decision Theory*. Dordrech: D. Reidel Publishing Co., 1984.

Allais, M. (1988a). The General Theory of Random Choices in Relation to the Invariant Cardinal Utility Function and the Specific Probability Function. The (U, θ) −Model: A General Overview. In: B.R. Munier, ed., *Risk, Decision and Rationality*. Dordrecht: D. Reidel Publishingt Co., 1988.

Allais, M. (1988b) Three Theorems on the Theory of Cardinal Utility and Random Choice. In: *Essays in Honour of Werner Leinfellner. Theory and Decision,* in press.

Anand, P. Are the Preference Axioms Really Rational? *Theory and Decision*, (Vol. 23, 1987), pp 189-214.

Bostic, R., Herrnstein, R.J. & Luce R.D. the Effect on the Preference-Reversal Phenomenon of Using choice Indifferences. Submitted, 1988.

Chew, S.H. a Generalization of the Quasilinear Mean with Applications to Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox. *Econometrica* (Vol. 51, 1983), pp 1065-1092.

Chew, S.H. & Epstein, L.G. (1987a). A Unifying Approach to Axiomatic Non-Expected Utility Theories, Manuscript, 1987.

Chew, S.H. & Epstein, L.G. (1987b). Generalized and Implicit Gini Indices. Manuscript, 1987.

Chew, S.H. Epstein, L.G. & Segal, U. Invariant Mean Values and Measures of Income Inequality. Manuscript, 1987.

Ellsberg, D. Risk, Ambiguity, and the Savage Axioms. *Quarterly Journal of Economics* (Vol. 75, 1961), pp 643-669.

Fishburn, P.C. Nontransitive Measurable Utility. *Journal of Mathematical Psychology* (Vol. 26, 1982),pp 31-67.

Fishburn, P.C. Transitive Measurable Utility. *Journal of Economic Theory* (Vol. 31, 1983), pp 293-317.

Gilboa, I. Expected Utility with Purely Subjective Non-Additive Probabilities. *Journal of Mathematical Economics* (Vol 16, 1987) pp 65-88.

Grether, D.M. & Plott, C.R. Economic Theory of Choice and the Preference Reversal Phenomenon. *The American Economic Review* (Vol. 69, 1979) pp 623-638.

Kahneman, D. & Tversky, A. Prospect Theory: An Analysis of Decision under Risk. *Econometrica* (Vol. 47, 1979), pp 263-291.

Karmarkar, U.S. Subjectively Weighted Utility: A Descriptive Extension of the Expected Utility Model. *Organizational Behavior and Human Performance* (Vol. 21, 1978), pp 61-72.

Lichtenstein, S. & Slovic, P. Reversals of Preference between Bids and Choices in Gambling Decisions. *Journal of Experimental Psychology* (Vol. 89, 1971), pp 46-55.

Loomes, G. & Sugden R. Regret Theory: An Alternative Theory of Rational Choice under Uncertainty. *Economic Journal* (Vol. 92, 1982), pp 805-824.

Loomes, G & Sugden, R. A Rationale for Preference Reversal. *American Economic Review* (Vol. 73, 1983), pp 428-432.

Luce, R.D. Uniqueness and Homogeneity of Ordered Relational Structures. *Journal of Mathematical Psychology* (Vol. 30, 1986), pp 391-415.

Luce, R.D. & Narens, L. Classification of Concatenation Measurement Structures According to Scale Type. *Journal of Mathematical Psychology* (Vol. 29, 1985), pp 1-72.

Luce, R.D. & Narens, L. Measurement, Theory of. In: J. Eatwell, M. Milgate, and P. Newman, eds *The New Palgrave: A Dictionary of Economic Theory and Doctrine*. New York: The Macmillan Press, 1987.

Machina, M.J. "Expected Utility" Analysis without the Independence Axiom. *Econometrica* (Vol. 50, 1982), pp 277-323.

Machina, M.J. Choice under Uncertainty: Problems Solved and Unsolved. *Economic Perspectives* (Vol. 1, 1987), pp 121-154.

Munera, H.A. The Generalized Means Model (GMM) for Non-Deterministic Decision Making: Its Normative and Descriptive Power, Including Sketch of the Representation Theorem. *Theory and Decision* (Vol. 18, 1985), pp 173-202.

Munera, H.A. The Generalized Means Model (GMM) for Non-Deterministic Decision Making: A Unified Treatment for the Two Contending Theories. In: L. Daboni et al, eds. *Recent Developments in the Foundation of Utility and Risk Theory*. Amsterdam: D. Reidel Publishing Co., 1986.

Munera, H.A. & de Neufville, R. A Decision Analysis Model when the Substitution Principle is Not Acceptable. In: B.P. Stigum and F. Wenstop, eds., *Foundations of Utility and Risk Theory with Applications*. Amsterdam: D. Reidel Publishiing Co., 1983

Pearson, T. Reversal of Preferences: Artifacts or Instransitivities. Manuscript, 1986.

Quiggin, J. A Theory of Anticipated Utility. *Journal of Economic Behavior and Organization* (Vol. 3, 1982), pp 323-343.

Röell, A. Risk Aversion in Quiggin and Yarri's Rank-Order Model of Choice under Uncertainty. *The Economic Journal* (Vol. 97, 1987), pp 143-159.

Segal, U (1987a). The Ellsberg Paradox and Risk Aversion: An Anticipated Utility Approach. *International Economic Review* (Vol. 28, 1987), pp 175-202.

Segal, U. (1987b). Some Remarks on Quiggin's Anticipated Utility. *Journal of Economic Behavior and Organization* (Vol. 8, 1987), pp 145-154.

Segal, U. (1987c). Axiomatic Representation of Expected Utility with Rank Dependent Probabilities. Manuscript, 1987.
Sen, A. *On Economic Inequality*. London: Oxford University Press, 1973.
Tversky, A & Kahnemann D. Rational Choice and the Framing of Decisions. *Journal of Business* (Vol. 59, 1986), pp S251-S278.
Tversky, A., Sattath, S. & Slovic, P. Contingent Weighting in Judgment and Choice. *Psychological Review*, in press, 1988.
Weber, M. & Camerer, C. Recent Developments in Modelling Preferences under Risk. *OR Spektrum* (Vol. 9, 1987), pp 129-151.
Yaari, M.E. The Dual Theory of Choice under Risk. *Econometrica* (Vol. 55, 1987), pp 95-115.

Appendix: Solution of a Linear System.

by Kenneth Manders, *University of Pittsburgh*

Let $n \geqslant 2$ be fixed integer, and let $\Pi(n)$ be the set of permutations of the set $\{i = 1, \ldots, n\}$. We regard permutations as ordered sequences; two permutations are *adjacent* if they are identical except for transpostition of two adjacent terms; these are said to be the *affected* terms, the others are said to be *unaffected*.

We first characterize the solution of the system of linear equations in unknowns $a_i(\rho)$, $i = 1, \ldots, n$; $\rho \in \Pi(n)$.

$$\sum_{i=1}^{n} a_i(\rho) = 1, \qquad \text{(each } \rho \in \Pi(n)\text{)}; \tag{A1}$$

$$a_i(\rho) = a_i(\rho'), \qquad \text{(for any two adjacent permutations } \rho \text{ and } \rho'\text{, and any unaffected } i\text{)}. \tag{A2}$$

It is convenient to introduce *reduced parameters* $b_i(\alpha)$, where $i = 1, \ldots, n$ and α is any subset (including the empty set) of $\{1, \ldots, n\} - \{i\}$. The reason for doing so is that (A2) intails the following much stronger constraint. For any $i \in \{1, \ldots, n\}$ and permutation ρ, let $\alpha(i, \rho)$ denote the (possibly empty) set of elements that are to the left of i under ρ. The constraint is that

$$a_i(\rho) = a_i(\rho'), \qquad \text{(for any permutations } \rho \text{ and } \rho' \text{ with } \alpha(i, \rho) = \alpha(i, \rho')\text{)}.$$

This follows because any two such permutations can be connected by a path of adjacent permutations, each pair leaving i unaffected.

Thus we can eliminate the original unknowns in favor of the reduced parameters $b_i(\alpha)$, $i = 1, \ldots, n$; α a subset of $\{1, \ldots, n\} - \{i\}$:

$$a_i(\rho) = b_i(\alpha(i, \rho)), \qquad \text{(each } i \text{ and } \rho\text{)}. \tag{A2'}$$

Claim. *Given the definition* (A2′), *the system* (A1), (A2) *is equivalent to the following conditions on the reduced parameters* $b_i(\alpha)$, $i = 1, \ldots, n$; α *a subset of* $\{1, \ldots, n\} - \{i\}$:

$$b_i(\alpha) + b_j(\alpha \cup \{i\}) = b_i(\alpha \cup \{j\}) + b_j(\alpha), \qquad (any\ \alpha \subseteq \{1, \ldots n\} - \{i,j\}); \tag{A1a′}$$

$$b_1(\emptyset) = 1 - (b_2(\{1\}) + \cdots + b_n(\{1, \ldots (n-1)\})). \tag{A1b′}$$

First, note that, given (A2), (A1) is equivalent to the two constraints

$$a_i(\rho) + a_j(\rho) = a_i(\rho') + a_j(\rho'), \qquad \text{(for any two adjacent permutations } \rho \text{ and } \rho'\text{, and the affected } i, j\text{);} \tag{A1a}$$

$$\sum_{i=1}^{n} a_i(\iota) = 1, \qquad \text{(for the identity permutation } \iota\text{).} \tag{A1b}$$

Clearly, both (A1a) and (A1b) are necessary in the presence of (A1) and (A2); conversely, any permutation may be connected to the identity permutation via a finite path of adjacent permutations, and (A1a) and (A2) then entail that the sum over each permutation is the same as the sum over the previous one.

Given (A2′), (A1b) is equivalent to

$$b_1(\phi) + b_2(\{1\}) + \cdots + b_n(\{1, \ldots (n-1)\}) = 1,$$

which is in turn equivalent to (A1b′). Consider two adjacent permutations ρ and ρ', with i,j affected. Without loss of generality, we may assume that $i < j$ and that i is to the left of j in ρ. Let $\alpha = \alpha(i, \rho)$ be the set of terms to the left of i in ρ. Given (A2′), we then obtain the corresponding instance of (A1a′) from (A1a). Any such combination α, i and j will arise from some instances of (A1a); thus given (A2′), (A1a) entails (A1a′). Conversely, any instance (A1a) correponds to some such α, i and j; and with (A2′) that instance of (A1a) follows from the corresponding instance of (A1a′).

In all, we conclude that given (A2′), the constraints (A1), (A2) are equivalent to (A1a′), (A1b′), as claimed. Thus the solutions to (A1), (A2) correspond exactly to those of (A1a′), (A1b′), via (A2′). We now come to our central claim.

Proposition. *The solutions to the system* (A1), (A2) *correspond uniquely to arbitrary choices of the* $2^n - 2$ *reduced parameters* $b_i(\alpha)$, α *any subset of* $\{1, \ldots, (i-1)\}$, $i = 2, \ldots, n$.

By the claim, it suffices to show the result for the solutions of (A1a′), (A1b′). Any solution of (A1a′), (A1b′) uniquely determines the reduced parameters $b_i(\alpha)$, α any subset of $\{1, \ldots, (i-1)\}$, $i = 2, \ldots, n$. It remains to show the converse. Thus we now suppose such a choice fixed, and show that it uniquely determines a solution to (A1a′), (A1b′). Of course, (A1b′) gives us the value of $b_1(\emptyset)$. ($\emptyset$ is the only subset of $\{1, \ldots, (i-1)\}$, $i = 1$.) It suffices to show that the given reduced parameters $b_i(\alpha)$, α

any subset of $\{1, \ldots, (i-1)\}$, $i = 1, \ldots, n$ (the case $i = 1$ now included by virtue of A1b′) uniquely determine all the remaining reduced parameters subject to (A1a′).

For each reduced parameter $b_i(\alpha)$, we define *the rank of* $b_i(\alpha)$ as the number of elements of α which exceed i. Thus the reduced parameters we have so far, $b_i(\alpha)$, α any subset of $\{1, \ldots, (i-1)\}$, $i = 1, \ldots, n$, are exactly those of rank 0. A look at the ranks in (A1a′), and recalling that $i < j$, shows that the third term has rank one higher than the greatest rank of the other three terms. Thus (A1a′) inductively constrains all reduced parameters of nonzero rank; we make this explicit by rearranging terms in (A1a′):

$$b_i(\alpha \cup \{j\}) = (b_i(\alpha) + b_j(\alpha \cup \{i\})) - b_j(\alpha), \qquad (\text{any } \alpha \subseteq \{1, \ldots n\} - \{i,j\}). \tag{A1a′}$$

We need to control this inductive definition more closely. Given $b_i(\alpha)$ not yet defined, such that all reduced parameters of smaller rank have already been defined, let $j = max(\alpha)$. Denote by α subscripted by one or more elements, such as j, the result of deleting the subscripted elements from α. We define $b_i(\alpha)$ by

$$b_i(\alpha) = (b_i(\alpha_j) + b_j(\alpha_j \cup \{i\})) - b_j(\alpha_j).$$

This uniquely determines each reduced parameter. It remains to show that the reduced parameters so defined satisfy (A1a′). Define the *rank* of an instance of (A1a′) as the rank of the maximal-rank term in that instance. We show the result by induction on the rank of instances of (A1a′). Therefore, suppose all instances of (A1a′) of rank less than m are satisfied. Each instance of (A1a′) of rank m has as maximal-rank term a reduced parameter $b_i(\alpha)$ of rank m. Let $j = max(\alpha)$. Then $b_i(\alpha)$ was defined by the instance of (A1a′) (shown just above) associated with α, i and j, which is therefore satisfied; it suffces to show that the instance associated with any other element k of α greater than i is also satisfied:

$$b_i(\alpha) = (b_i(\alpha_k) + b_k(\alpha_k \cup \{i\})) - b_k(\alpha_k).$$

Thus we must show these two expressions for $b_i(\alpha)$ equal. Evaluating some terms of lower rank in these expressions using (A1a′), we find

$$b_i(\alpha_j) = (b_i(\alpha_{jk}) + b_k(\alpha_{jk} \cup \{i\})) - b_k(\alpha_{jk});$$

$$b_j(\alpha_j \cup \{i\}) = (b_j(\alpha_{jk}) + b_k(\alpha_k \cup \{i\})) - b_k(\alpha_{jk} \cup \{i\});$$

$$b_j(\alpha_j) = (b_j(\alpha_{jk}) + b_k(\alpha_{jk} \cup \{j\})) - b_k(\alpha_{jk});$$

$$b_i(\alpha_k) = (b_i(\alpha_{jk}) + b_j(\alpha_{jk} \cup \{i\})) - b_j(\alpha_{jk}).$$

Substituting these expressions in the two original expressions, and noting that $\alpha_{jk} \cup \{j\} = \alpha_k$, we find that the original expressions are indeed equal. By induction on m, we conclude that the system of reduced parameters constructed does satisfy (A1a'). This completes the proof of the proposition.

Corollary. *For any $n \geqslant 2$, the dimension (equivalently, the number of free parameters) of the solution set of the following system of linear inequalities, in unknowns $a_i(\rho)$, $i = 1, \ldots, n$; $\rho \in \Pi(n)$, is $2^n - 2$.*

$$\sum_{i=1}^{n} a_i(\rho) = 1, \qquad (\rho \in \Pi(n)); \tag{A1}$$

$$a_i(\rho) = a_i(\rho'), \qquad \textit{(for any two adjacent permutations } \rho \textit{ and } \rho' \textit{, and any unaffected } i\textit{)}; \tag{A2}$$

$$a_i(\rho) \geqslant 0, \qquad (i = 1, \ldots, n;\ \rho \in \Pi(n)). \tag{A3}$$

In view of the above, it suffices to show that the solution set of (A1)-(A2) assumes its full dimensionality within the positive cone. Now at least the constant solution ($a_i(\rho) = 1/n$, each i, ρ) lies strictly within the positive cone (A3). The solution set of (A1)-(A2) is convex, because the system is linear. Thus the positive cone contains a neighborhood of the constant solution in the solution set; but any neighborhood in the solution set is $(2^n - 2)$-dimensional (and indeed, parameterized by the reduced parameters specified in the proposition above).

Acknowledgments

Work carried out while the author was a Fellow of the Center for Advanced Studies in the Behavioral Sciences, and a National Endowment for the Humanities Fellow. I am grateful for financial support provided by NSF Grant SES 85-11229, by NEH Grant #FC-20060-85 and the Andrew W. Mellon Foundation.