Journal of Risk and Uncertainty, 4:29–59 (1991)

Rank- and Sign-Dependent Linear Utility Models for Finite First-Order Gambles

R. DUNCAN LUCE
Irvine Research Unit in Mathematical Behavioral Sciences, Social Science Tower, University of California, Irvine, CA 92717

PETER C. FISHBURN
AT&T Bell Laboratories, Murray Hill, NJ 07974-2070

Key words: generalized SEU, generalized prospect theory, rank-dependent utility, sign-dependent utility

Abstract

Finite first-order gambles are axiomatized. The representation combines features of prospect and rank-dependent theories. What is novel are distinctions between gains and losses and the inclusion of a binary operation of joint receipt. In addition to many of the usual structural and rationality axioms, joint receipt forms an ordered concatenation structure with special features for gains and losses. Pfanzagl's (1959) consistency principle is assumed for gains and losses separately. The nonrational assumption is that a gamble of gains and losses is indifferent to the joint receipt of its gains pitted against the status quo and of its losses against the status quo.

This article extends four previous papers (Luce, 1988, in press (a); Luce and Narens, 1985; Kahneman and Tversky, 1979) that have developed partial, more-or-less descriptive theories for choices among uncertain alternatives or, briefly, gambles. When the probabilities of the events are known, we speak of risk and will explicitly modify the word *gamble*.

For binary gambles with known probabilities and for trinary ones in which one consequence is no change from the status quo, Kahneman and Tversky (1979) developed a weighted average theory in which the weights depended both on the rank order of the consequences and on their relation to no change. Aside from Edward's (1962) discussion of the importance of the status quo and the fact that nonadditive weights may entail ratio-scale representations, theirs is the first well-developed ratio-scale theory of utility of which we are aware.[1] They called it *prospect theory*. Subsequently, Fishburn (1982, 1988) has worked out nontransitive ratio theories.

Reprints may be obtained from either author. Luce's work was supported, in part, by the National Science Foundation grant IRI-8996149 to the University of California, Irvine. We thank an anonymous referee and the editor for helpful suggestions, and Paul Slovic for bringing to our attention his 1968 paper with S. Lichtenstein.

For binary gambles and indefinite iterations of them, Luce and Narens (1985) axiomatized a pure *rank-dependent*[2] (RD) theory without any assumption that event probabilities are known. Luce (1988) extended this to a rank-dependent theory for arbitrary gambles with finitely many distinct consequences and discussed its relation to earlier rank-dependent theories. For the most part, the earlier theories were less general because they considered only choices between money gambles with known probability distributions, i.e., between random variables. More importantly, treating gambles as random variables systematically ignores the structures within multistage gambles. This is fine for applications to one-stage gambles, so long as it is recognized that when one assumes that decision problems in extensive form can be replaced by their equivalent normal form, then it becomes very difficult to avoid subjective expected utility (SEU) except by introducing violations of monotonicity (sometimes called cancellation). But, as pointed out in Luce (in press (a), in press (b)), existing data give very little reason to reject monotonicity in the consequences, be these either elementary ones or imbedded gambles. Far more suspect is the assumption of indifference between a multi-stage gamble and its formally equivalent single-stage reduction, which is sometimes called *reduction of compound lotteries* or an *accounting equivalence*.

One must be careful not to misinterpret these remarks. They point out that an assumption of the accounting equivalences goes a long way toward forcing SEU, and so to the extent that SEU is wrong, as shown by the Allais paradox, the accounting equivalences may well be implicated as wrong rather than the monotonicity of the consequences. It does not follow, one way or the other, from evidence against SEU that is based entirely on single-stage gambles that monotonicity of the consequences is necessarily implicated as the source of the problem. That issue is testable, and the data to date favor monotonicity of the consequences. The difficulties with SEU appear to lie elsewhere.

The last paper of the above-mentioned quartet (Luce, in press (a)) examined and axiomatized a *binary rank- and sign-dependent* (RSD) theory in which the weights depend on both the relation of the consequence to no change in the status quo (sign) and to the other consequence (rank). It was shown that this ratio-scale theory is a natural generalization of three more special theories: the binary rank-dependent, the binary sign-dependent, and the binary prospect theory.

The most unusual feature of the last axiomatization was its use of an operation $\oplus$ in addition to the usual mixing operations $\bigcirc_E$. The symbol $a \bigcirc_E b$ means that the consequence is a if event E occurs and b otherwise, where a and b may be pure consequences, such as sums of money, or other gambles. The symbol $a \oplus b$ is interpreted as meaning that both consequences, a and b, are received. One basic assumption of that theory is that utility is additive over $\oplus$, i.e., $U(a \oplus b) = U(a) + U(b)$. Under reasonable assumptions about money outcomes, this was shown to imply that U is a pair of power functions of money, with different constants for consequences above and below the status quo. This reduces to proportionality in monetary amount if, for sums of money x and y, $x \oplus y = x + y$. The major significance in having $\oplus$ is that it provides a formal way to describe the sort of editing of gambles that was discussed but not formalized in prospect theory.

The first aim of the present article is to axiomatize the joint receipt operation so as to have the following order-preserving numerical representation U (theorem 2): Let e denote the status quo and let g and h be any two gambles. Then

$$U(g \oplus h) = \begin{cases} A(+)U(g) + B(+)U(h) + C(+)U(g)U(h), & \text{if } g \gtrsim e, h \gtrsim e, \\ A(+)U(g) + B(-)U(h), & \text{if } g \gtrsim e \gtrsim h, \\ A(-)U(g) + B(+)U(h), & \text{if } h \gtrsim e \gtrsim g, \\ A(-)U(g) + B(-)U(h) + C(-)U(g)U(h), & \text{if } e \gtrsim g, e \gtrsim h; \end{cases}$$

$$U(e) = 0.$$

An interesting reason underlies this generalization of pure additivity. Theorem 1 establishes that the nonadditive form on either side of the status quo is pretty much dictated by a weighted-average representation for gambles and a rational-consistency axiom of Pfanzagl (1959) for $\oplus$. The additive form for the joint receipt of gains and losses is to a large extent dictated by the monotonicity of $\oplus$.

One result of this more general form is that if $C \neq 0$, then the utility of money is an exponential of a power function of money (see equation (5)). Such functions are unusual in that an initially diminishing marginal utility gives way ultimately to an increasing one.

Second, we take further advantage of the fact that having $\oplus$ provides a simple way to describe formally an editing process designed to simplify gambles. Such editing was discussed informally by Kahneman and Tversky (1979). Specifically, we assume that a gamble involving only gains is judged indifferent to the joint receipt of the smallest gain and of the gamble that results from "subtracting" that consequence from each of the gains. A similar assumption is made for gambles involving only losses. The axiom (R3) underlying this editing is a slightly weakened version of Pfanzagl's (1959) consistency principle, and it appears to be an additional form of rationality beyond the usual ones for gambles. This subtraction procedure is used inductively, and it results in a RD representation for the domain of gains and a distinct RD one for the domain of losses. In sharp contrast to the RD theory proposed earlier, this theory makes quite transparent why rank dependence arises naturally as a result of editing gambles for which only gains or only losses can occur.

Third, the case of mixed gains and losses obviously must be addressed. The assumption made, which is the single nonrational one in the theory, is that the decision maker decomposes such a gamble into a joint receipt of two closely related gambles that, together, are treated as indifferent to the given one. The one involves just the gains with the losses replaced by the status quo; the other, just the losses with the gains replaced by the status quo. The axiom (D1) embodies the maxim: evaluate the possible gains, evaluate the possible losses, and then base your decision on how those two balance out. Although such advice clearly is not rational, to many it seems plausible, and some data support its descriptive accuracy (Slovic and Lichtenstein, 1968).

In brief summary, a representation of the following character is axiomatized. Let a first-order gamble, g, be a mapping from a finite partition of an event E into a set of consequences. Number the events of the partition from best to worst according to the preference ranking of the associated consequences. Let $E(+)$ denote the union of events whose consequences are gains (i.e., positive relative to the status quo), and let $E(-)$ be the union of events whose consequences are losses (i.e., negative relative to the status quo). The representation is that there exists a ratio scale U over uncertain alternatives and pure consequences and unique weights over events of the form $S^i[E(i) \mid E]$, $i = +, -$, where $+$ is used for gains and $-$ for losses, such that $U(g)$ consists of two rank-dependent utility expressions over the events $E(i)$ that are weighted by $S^i[E(i) \mid E]$.

This theory treats gambles with both gains and losses in a manner similar to, but more general than, prospect theory, whereas alternatives that have consequences of only one sign act in a purely rank-dependent fashion, similar to, but more general than, Luce's (1988) earlier theory. The result is a significant generalization of prospect theory in that this theory applies to gambles based on uncertain events, not just to the case of risk, and to any finite, first-order gamble, not just to those with a single gain and a single loss.

Four types of axioms are to be distinguished: structural, rational in the traditional sense of gambles (especially transitivity and monotonicity of consequences, but not events), rational in a sense that is appropriate to joint receipt (including monotonicity and a limited form of consistency), and a decomposition axiom that conceivably may be descriptive and may seem plausible, but certainly is not rational. Because the consistency and decomposition conditions involving joint receipt are both crucial to the representation derived and relatively little studied empirically, they should become the focus of empirical attention.

1. Notation

It is essential to have a notation that allows us to keep track of what we mean by certain derived gambles. We begin with two structures: $\mathscr{E}$ is an algebra of events, and $\mathscr{C}$ is a set of pure consequences, such as money or consumer items, but not gambles. Within $\mathscr{C}$, there is a special null consequence e that represents no change from the status quo.

For any finite partition[3] $\{E_j\}$ of $E \in \mathscr{E}$, a function from $\{E_j\}$ into $\mathscr{C}$ is called a *first-order (finite) gamble*. Denote the family of all first-order, finite gambles by $\mathscr{g}_1$. Any function from a finite partition into $\mathscr{g}_1 \cup \mathscr{C}$ with at least one value in $\mathscr{g}_1$ is called a second-order (finite) gamble. Let $\mathscr{g}$ consist of all the first-order gambles together with $\mathscr{C}$ and at least all of the second-order gambles characterized by the axioms (in particular, axiom S2).

$\succsim$ is a binary relation on $\mathscr{g}$ that is interpreted as a preference relation over gambles. The converse of $\succsim$ is denoted by $\precsim$. We define $\sim$ and $\succ$ in the usual ways: $g \sim h$ iff both $g \succsim h$ and $g \precsim h$; $g \succ h$ iff $g \succsim h$ and not $(g \precsim h)$.

$\oplus$ is a binary operation on $\mathscr{g}$, which is interpreted as joint receipt.

So the entire structure is $\mathscr{P} = \langle \mathscr{E}, \mathscr{C}, e, \mathscr{g}, \succsim, \oplus \rangle$. The task is to axiomatize it in such a way that it leads to a plausible descriptive theory of utility. That task is divided into two distinct subtasks. The first is to axiomatize the representation of $\oplus$ without any real regard to the internal structure of the gambles. That done, we axiomatize the gambles, taking advantage of our knowledge about the representation of $\oplus$.

2. Axiomatization of joint receipt

2.1. *Proposed representation*

The axiomatization of $\langle \mathcal{G}, \succsim, \oplus \rangle$ to be described is aimed at justifying the following real representation $U: \forall g, h \in \mathcal{G}$,

$$g \succsim h \text{ iff } U(g) \geq U(h), \tag{1a}$$

$$U(g \oplus h) = \begin{cases} A(+)U(g) + B(+)U(h) + C(+)U(g)U(h), & \text{if } g \succsim e, h \succsim e, \\ A(+)U(g) + B(-)U(h), & \text{if } g \succsim e \succsim h, \\ A(-)U(g) + B(+)U(h), & \text{if } h \succsim e \succsim g, \\ A(-)U(g) + B(-)U(h) + C(-)U(g)U(h), & \text{if } e \succsim g, e \succsim h; \end{cases} \tag{1b}$$

$$U(e) = 0, \tag{1c}$$

where the weights A and B are positive, $C(+) \geq 0$, and $C(-) \leq 0$.

The form given for the positive and negative quadrants is formally similar to a multilinear form from multiattribute utility theory that uses $A(j) = B(j) = 1$ (Keeney and Raiffa, 1976, p. 250). Theorem 1 below gives a moderately compelling reason for looking at representations having the slightly more general character of equation (1b). It is rather less clear what to assume for the mixed cases, and the form chosen strikes us as a reasonable compromise between the simplest additive form, i.e., $U(g) + U(h)$, and more complex possibilities. Among its features is that U is continuous at e and monotonic throughout, and the form leads to a natural generalization of prospect theory. It is not difficult to show that were one to try to introduce a product term $CU(g)U(h)$ in the mixed case, monotonicity forces $C = 0$.

The constants $C(i)$ have the dimension $1/U$, and for that reason the representation is a ratio scale. Unlike many conventional measurement structures, the admissible ratio-scale transformations of multiplication by positive constants do not correspond to automorphisms of the structure.

2.2. *Motivation for the representation*

Our first result, which is concerned only with binary gambles, presupposes (which is formulated as axiom J1 later) that the structure involved has a real representation U with $U(e) = 0$. The main purpose of the result is to show that with a certain degree of rationality holding in either the domain of gains or losses, the form of the utility function over $\oplus$ given in equation (1b) is very close to there being a rank-dependent representation for binary gambles. Moreover, these two properties for the utility function imply a major postulate of rationality.

In an abuse of notation, we continue to write $\oplus$ for the numerical operation corresponding to the qualitative one. The binary gambles are written in operator notion. In stating the result, we assume that $\oplus$ is *accumulative* in the sense that if min $\{x, y\} \geq 0$, then $x \oplus y \geq \max\{x, y\}$ and if $\max\{x, y\} \leq 0$, then $x \oplus y \leq \min\{x, y\}$.

A binary gamble based on the universal event of $\mathscr{E}$ is denoted in operator fashion: $x \otimes_E y$, meaning that the consequence is x if E occurs and y otherwise.

Theorem 1. Suppose $\mathscr{E}$ is a family of events and $\langle Re^j \cup \{0\}, \geq, \oplus, \otimes_E\rangle_{E\in\mathscr{E}}, j = +, -$, is such that

1. $\oplus$ and $\otimes_E$ are closed, strictly increasing, binary operations; $\oplus$ is accumulative; and $0 \oplus 0 = 0$;
2. U is a strictly increasing map from Re onto Re with $U(0) = 0$ such that two of (i), (ii), and (iii) obtain:

(i) (a) $\forall x, y \in Re^j, \exists A(j), B(j) \in Re^+$ and $\exists C(j) \in Re^j$ with $C(j)$ having the dimension of $1/Uj$ such that

$$U(x \oplus y) = A(j)U(x) + B(j)U(y) + C(j)U(x)U(y); \tag{2}$$

(b) 0 is a generalized zero in the sense that both $U(x \otimes_E 0)$ and $U(0 \otimes_E x)$ are proportional to $U(x)$, and the constants of proportionality as a function of E span (0, 1).

(ii) (a) $(\forall x, y \in Re^j, \forall E \in \mathscr{E}: \exists S_i^j(E) \in (0, 1), i = >, <)$

$$U(x \otimes_E y) = \begin{cases} S_>^j(E)U(x) + [1 - S_>^j(E)]U(y), & \text{if } x \gtrsim y \\ S_<^j(E)U(x) + [1 - S_<^j(E)]U(y), & \text{if } x < y; \end{cases} \tag{3}$$

(b) the S_i^j span (0, 1).

(iii) (Distribution) $(\forall x, y \in Re^j)$,

$$(x \otimes_E 0) \oplus y = (x \oplus y) \otimes_E (0 \oplus y) \text{ and}$$
$$y \oplus (0 \otimes_E x) = (y \oplus 0) \otimes_E (y \oplus x).$$

Then the third one obtains.

Corollary. In case either $A \neq 1$ or $B \neq 1$ in (i)(a), then (i)(b) can be replaced by the weaker assumption that the values of the functions relating $U(x \otimes_E 0)$ and $U(0 \otimes_E x)$ to $U(x)$ span (0, 1) at $U(x) = 1$.

Proofs of theorem 1 and all ensuing results will be found in the appendix.

The major rationality postulates of theorem 1 are the monotonicity of assumption (1) and the distribution property (iii). The latter says that the joint receipt of a gamble and a fixed sum y is viewed as indifferent to a gamble in which each outcome is the joint receipt of the original outcome jointly with y. The significance of this theorem is that in the presence of such rationality, rank dependence and the form of equation (1b) in the domains of gains and losses are substantially equivalent. As noted earlier, this leaves open the mixed case, which was dictated by less formal considerations.

2.3. *Axioms giving rise to the representation*

The following seven axioms are sufficient for the representation of equation (1).

Axiom J1. $\langle \mathcal{G}, \succsim \rangle$ is order-dense, Dedekind complete,[4] and has $g \succ e \succ h$ for some $g, h \in \mathcal{G}$.

Axiom J2. $\oplus$ is closed, $e \oplus e \sim e$, and is monotonic in the sense that $\forall g, h, k \in \mathcal{G}$,

(i) if $g \sim h$, then $g \oplus k \sim h \oplus k$ and $k \oplus g \sim k \oplus h$, and
(ii) if $g \succ h$, then $g \oplus k \succ h \oplus k$ and $k \oplus g \succ k \oplus h$.

Axiom J3. For any three of the four terms in any one of

$g \oplus h \sim k \oplus l$, $(g \oplus h) \oplus k \sim l$, and $k \oplus (g \oplus h) \sim l$,

the fourth terms exists so that $\sim$ holds.

Although neither axiom J1 nor the solution-of-equations axiom J3 is necessary for the representation, axiom J2 is easily seen to be. Necessity for (ii) uses the positivity of the A and B weights. Note also that axiom J1 presumes the necessary condition that $\succsim$ is a weak order, i.e., transitive and complete. Transitivity combines with the monotonicity conditions (i) and (ii) of axiom J2 to give $g \oplus k \succ h \oplus l$ whenever $g \succsim h, k \succsim l$ and at least one $\succsim$ is strict.

For the remaining axioms, $\mathcal{G}$ is partitioned into its nonnegative and nonpositive parts:

$$\mathcal{G}^+ = \{g \mid g \succsim e\} \text{ and } \mathcal{G}^- = \{k \mid k \precsim e\}.$$

Axiom J4. The substructures $\langle \mathcal{G}^+ \times \mathcal{G}^-, \succsim \rangle$ and $\langle \mathcal{G}^- \times \mathcal{G}^+, \succsim \rangle$, in which (g, h) is defined as $g \oplus h$, are both additive conjoint structures.

Axiom J5. $\forall f, g, h \in \mathcal{G}^+, \forall, j, k, l \in \mathcal{G}^-$,

(i) if $f \oplus e \sim g \oplus j$, $e \oplus f \sim k \oplus g$, and $h \oplus e \sim f \oplus j$, then $e \oplus h \sim k \oplus f$, and
(ii) if $j \oplus e \sim k \oplus f$, $e \oplus j \sim g \oplus k$, and $l \oplus e \sim j \oplus f$, then $e \oplus l \sim g \oplus j$.

Axiom J6. $\forall g, h \in \mathcal{G}^+, \forall k, l \in \mathcal{G}^-$, any two of the following implies the third:

$$g \oplus e \sim e \oplus h, e \oplus k \sim l \oplus e, g \oplus k \sim l \oplus h.$$

Axiom J7. $\forall g, h, i, j, k \in \mathcal{G}^+, \forall l, m, n, p, q \in \mathcal{G}^-$,

(i) if $h \oplus e \sim g \oplus l, j \oplus e \sim i \oplus l$, and $(h \oplus k) \oplus e \sim (g \oplus k) \oplus m$, then $(j \oplus k) \oplus e \sim (i \oplus k) \oplus m$;
(ii) if $m \oplus e \sim l \oplus g, p \oplus e \sim n \oplus g$, and $(m \oplus q) \oplus e \sim (l \oplus q) \oplus h$, then $(p \oplus q) \oplus e \sim (n \oplus q) \oplus h$.

The two expressions obtained by commuting every $\oplus$ term in (i) and (ii) also hold.

Axioms J4–J7 are necessary for the representation, as is easily verified with substitutions from equation (1). Axiom J4 obviously lies behind the middle two lines of equation (1b), but, as shown in our sufficiency proof, the tradeoff conditions of axioms J5–J7 are also needed there. These axioms are far from transparent, but we have not been able to simplify them in the absence of commutativity ($g \oplus h = h \oplus g$) and the true zero condition $g \oplus e = e \oplus g = g$. If the latter two conditions are assumed, then axioms J5–J7 can be replaced by the simpler

Axiom J7′. $\forall f, g, h \in \mathcal{G}^+, \forall j, k, l \in \mathcal{G}^-$,

(i) if $f \oplus j, g \oplus j \in \mathcal{G}^+$ and $(f \oplus j) \oplus h \sim (f \oplus h) \oplus l$, then $(g \oplus j) \oplus h \sim (g \oplus h) \oplus l$;
(ii) if $j \oplus f, k \oplus f \in \mathcal{G}^-$ and $(j \oplus f) \oplus l \sim (j \oplus l) \oplus h$, then $(k \oplus f) \oplus l \sim (k \oplus l) \oplus h$.

The effect of these changes on the representation is to force each A and B constant to equal 1.

Theorem 2. If $\langle \mathcal{G}, \succsim, \oplus, e \rangle$ satisfies axioms J1–J7, then it has the representation U of the form of equation (1) with U onto Re.

Corollary. Under the hypotheses of theorem 2,

(i) $A(j) > 0, B(j) > 0, j = +, -; C(+) \geq 0, C(-) \leq 0$;
(ii) $A(j) = B(j)$ iff $\oplus$ is commutative;
(iii) e is a zero of $\oplus$ ($\forall g \in \mathcal{G}, e \oplus g \sim g \oplus e \sim g$) iff $A(j) = B(j) = 1$, in which case $\oplus$ is commutative everywhere and associative within each of $\mathcal{G}^+$ and $\mathcal{G}^-$.

For the specialization $A(j) = B(j) = 1$ of the representation of equation (1b), $\oplus$ is commutative but not associative. Yet within each of the four quadrants, $\oplus$ is purely extensive (additive) in character, since it is easy to see that in the positive and negative domains, the transformation $V(g) = \log[1 + C(i)U(g)]$ satisfies $V(g \oplus h) = V(g) + V(h)$ (see the proof of theorem 3), and it has been assumed to be additive in the mixed positive–negative quadrants. Two different functions are needed to produce additivity in the same-sign and opposite-sign quadrants.

2.4. *Utility for money*

It is interesting to arrive at the form of U for money. We do so under two additional assumptions: that money has a unit structure representation,[5] i.e.,

$$\forall x, y \in Re^i, x \oplus y = yF_i(x/y), i = +, -,$$

and that in the utility representation of equation (1b), $A(j) = B(j) = 1$.

Theorem 3. Assume the conditions of theorem 2. If the underlying operation over the positive and over the negative real numbers both form unit structures, e is a zero of $\oplus$, and $\oplus$ is commutative, then the function U of equation (1) is of the following form: there exist positive constants $k(j)$ and $\beta(j), j = +, -$, such that for $x \in Re$:
If $C(j) = 0$

$$U(x) = \begin{cases} k(+)x^{\beta(+)}, & \text{if } x > 0, \\ -k(-)(-x)^{\beta(-)}, & \text{if } x < 0. \end{cases} \tag{4}$$

If $C(j) \neq 0$,

$$U(x) = \begin{cases} (\exp[k(+)x^{\beta(+)}] - 1)/C(+), & \text{if } x > 0, \\ (\exp[k(-)(-x)^{\beta(-)}] - 1)/C(-), & \text{if } x < 0. \end{cases} \tag{5}$$

Thus, the theory admits several qualitatively quite different utility functions for money. With $C = 0$, the possibilities are an increasing slope ($\beta > 1$), a constant slope ($\beta = 1$), or a decreasing slope ($\beta < 1$). With $C \neq 0$, the two possibilities are either an increasing slope ($\beta \geq 1$) or a slope that first decreases and then increases ($\beta < 1$). The switch-over in the latter case occurs when $x = [(1 - \beta)/k\beta]^{1/\beta}$. It is generally felt that increasing slopes do not accord with data. However, a person of the latter type is different in that s/he acts "normally" for modest sums but becomes insatiable for very large sums. This behavior seems to describe at least some very successful financial types. Whether it in fact holds for the rest of us remains untested and in most cases untestable.

We do not have any results about the form of U in the general case of theorem 2.

3. Axioms for choices between gambles

3.1. *Structural axioms*

In addition to axioms J1 and J3, which are structural, two more structural assumptions are used to ensure that the domain of consequences and gambles is sufficiently rich. The first of these requires the space of consequences itself to be very rich, in some sense as rich as the space of gambles. Certainly, this is not unreasonable if the set of consequences includes money.

Axiom S1. For each $g \in \mathcal{G}$ there exists a $c \in \mathcal{C}$ such that $c \sim g$.

To formulate our last structural axiom as well as a later rationality axiom, we need some additional notation. Suppose g is a first-order gamble based on the partition $\{E_j\}$ of an event E. Define the following three sets of events:

$$E(+) = \{\cup E_j \mid g(E_j) > e\},$$
$$E(0) = \{\cup E_j \mid g(E_j) \sim e\},$$
$$E(-) = \{\cup E_j \mid g(E_j) < e\}.$$

With these as domains, define gambles g^i for $i = +, 0, -$:

$$g^i(E_j) = \begin{cases} g(E_j), & \text{if } E_j \subseteq E(i) \\ \text{not defined otherwise.} \end{cases} \tag{6}$$

Now define the second-order gamble g_2 on the partition $\{E(+), E(0), E(-)\}$ of E:

$$g_2[E(i)] = g^i, i = +, 0, -.$$

We refer to g_2 as a *sign-partitioned* gamble, since it partitions g into its positive, null, and negative components.

Axiom S2. If g is a first-order gamble on the event E and g_2 is its second-order, sign-partitioned equivalent that is based on the partition $\{E(+), E(0), E(-)\}$, then $g_2 \in \mathcal{G}$.

3.2. *Rationality axioms*

Previous discussions of the data (Luce, in press (a), in press (b)) plus well-known normative considerations suggest that we retain the assumptions of both the transitivity of preferences and monotonicity of consequences for at least the binary gambles as being probably descriptive as well as rational.

The monotonicity axiom, which simply says that binary gambles, as mathematical functions into the space of consequences, are strictly monotonic increasing, is potentially controversial. As was pointed out in Luce (in press (a), in press (b)), so long as one does not automatically assume that indifference holds between a higher-order (or extensive) gamble and its corresponding, formally equivalent, first-order (normal) form, then there is no known descriptive problem in assuming monotonicity. When confronted with monotonicity that is presented in a transparent, structured gamble, such as explicitly replacing one consequence by a more preferred one, most people abide by it; only by combining monotonicity with a reduction to first-order form, as in the Allais paradox, does trouble arise. Since we do not assume all such reductions, those data seem largely irrelevant, and what data there are seem to support monotonicity (Kahneman and Tversky, 1979; Keller, 1985).

Axiom R1. As a function from an event partition into $\mathscr{C}$, a binary gamble is strictly increasing in each argument. In particular, if g and h are two gambles on the same partition $\{E_1, E_2\}$ that agree on one subevent but not on the other, say E_k, then

$$g(E_k) \gtrsim h(E_k) \text{ iff } g \gtrsim h.$$

Observe that any first-order gamble g is formally identical, in the sense that each consequence arises under exactly the same conditions, to the sign-partitioned second-order gamble g_2 that was defined just prior to axiom S2. For suppose c is some consequence that arises in g, and suppose it arises when the subevent E_j occurs, i.e., $g(E_j) = c$. If $c > e$, then $E_j \subseteq E(+)$, and so by the definition of g_2 it also arises if and only if E_j occurs. The argument is similar in the other two cases. So our next axiom is an accounting equivalence that asserts that the decision maker is indifferent between these two, formally equivalent formulations. This rational axiom seems to be moderately transparent, and such partitionings tend to arise quite spontaneously when subjects confront mixed gambles. Clearly it needs to be studied empirically in detail.

Axiom R2. If g is a first-order gamble on the event E and g_2 is its formally equivalent, second-order, sign-partitioned version, then

$$g \sim g_2.$$

Our next assumption is a form of distribution relating gambles to the operation $\oplus$; the assumption is essentially the consistency axiom of Pfanzagl (1959) applied to two of the four quadrants.

Axiom R3. Suppose that g is a first-order gamble on the partition $\{E_j\}$ with only positive (negative) or null consequences and that $c \in \mathscr{C}$ is a positive (negative) consequence. Define the gambles g_c and ${}_cg$ on $\{E_j\}$ by

$$g_c(E_j) = g(E_j) \oplus c \text{ and } {}_cg(E_j) = c \oplus g(E_j).$$

Then

$$g \oplus c \sim g_c \text{ and } c \oplus g \sim {}_cg.$$

Observe that axiom R3 is a generalization of the distribution property of theorem 1.

3.3. *Decomposition axiom*

Our final task is to explore what may happen to preferences over gambles with mixed consequences. The basic hypothesis is that second-order, sign-partitioned gambles of the sort discussed prior to axiom S2 can be decomposed into simpler second-order gambles using the operation $\oplus$ of joint receipt. These simpler gambles are also sign partitioned, but they involve a partition into only two rather than three subevents. For such binary gambles, it is convenient to use the operator notation of theorem 1.

Axiom D1. If g_2 is a signed-partitioned gamble of the form described prior to axiom S2, then for independent realizations of $E(+)$ and $E(-)$

$$g_2 \sim (g^+ \otimes_{E(+)} e) \oplus (g^- \otimes_{E(-)} e).$$

This postulate says that a sign-partitioned gamble is decomposed by the decision maker into the joint receipt of two sign-partitioned gambles, one of which involves only the positive part pitted against no change in the status quo and the other only the negative part pitted against no change. In form, it is much like a rational accounting equivalence, except that no one would hold that it is really rational, since the possibilities on the left are only single consequences whereas various combinations of them are possible on the right. Axiom D1 is a key assumption of the theory, one that deserves careful empirical examination. Basically it says that the major point of nonrationality on the part of decision makers concerns gambles with mixed positive and negative consequences.

For the binary case, the empirical literature refers to the right side of the expression in axiom D1 as a duplex gamble; see, for example, Slovic (1967). He spoke of the right side as being "... as faithful an abstraction of real life decision situations ... " as is the more conventional left side, but he did not make clear whether or not he believed that people treat them as equivalent, as we are assuming. Subsequently, Slovic and Lichtenstein (1968) carried out an empirical study of the binary version of axiom D1. In their experiment I, each of 19 subjects evaluated each standard gamble g_2 twice and each duplex gamble once by reporting minimum selling prices. Taking the variability in the replications as a measure of the "noise level," they found that " ... for a good majority of the [subjects], the differences between parallel duplex and standard bets were no larger than the differences between replications of the same standard bets ... ." This study needs replication using either choice indifference points or direct comparisons between the gamble and the joint-receipt pair. Moreover, a new study should include gambles with more than one gain and one loss.

4. Representations

4.1. *Preliminary results*

The main result to be demonstrated is that this mixture of rational and plausible assumptions leads to a fairly simple rank- and sign-dependent representation of utility. To some degree it is a substantial generalization of previous work. At the same time, it is distinctive in that the representation makes highly explicit the source of rank dependence, which has not really been true of previous rank-dependent theories. Toward that end, we first prove the following.

Theorem 4. Suppose axioms J1-J7, S1, and R1 hold.

(i) $\forall b, c \in \mathscr{C}$ with $b > c \gtrsim e$, there exist $c', c'', d', d'' \in \mathscr{g}^+ \cap \mathscr{C}$ such that $c \sim e \oplus c' \sim c'' \oplus e$ and $b \sim d' \oplus c' \sim c'' \oplus d''$. We write $c \ominus c' \sim e \sim c \ominus' c''$, $b \ominus c' \sim d'$, and $b \ominus'' c'' \sim d''$. A similar statement holds for $b < c \lesssim e$.

(ii) Suppose axiom R3 also holds, that g is a first-order gamble on the partition $\{E_j\}$ with only positive consequences, and $c \in \mathscr{C}$ is the minimum consequence of g. Define c' and c'' as in (i), and define first-order gambles g' and g'' by

$$g'(E_j) = g(E_j) \ominus c' \text{ and } g''(E_j) = g(E_j) \ominus' c''.$$

Then

$$g' \oplus c' \sim g \sim c'' \oplus g''.$$

A similar statement holds for any first-order gamble with only negative consequences, using the maximum of these negative consequences rather than the minimum.

The second assertion of the theorem simply says that any first-order gamble all of whose consequences have the same sign may be edited as follows. For the consequence that is nearest in preference to no change in the status quo, one finds what must be "subtracted" from it to be equivalent to the status quo, and then "subtracts" that amount from each of the other consequences; then that common amount is jointly received along with the modified gamble. In the noncommutative case, there are two ways that this can be done. In the simplest case, when e is a zero,[6] in the sense that for all a, $a \oplus e \sim e \oplus a \sim a$, then $b \sim b' \sim c$, and so there is only one form of subtraction.

Such isolation of a common factor is certainly a plausible reframing of a gamble, one that was used by Kahneman and Tversky (1979), and to some degree it seems highly rational. After all, if E_j occurs, one receives $g(E_j)$ under g and $[g(E_j) \ominus c] \oplus c \sim g(E_j)$ in the edited case.

It is unclear how one should carry out such editing for a gamble with both gains and losses. What should be taken as the common factor to be subtracted? Because of this ambiguity, the axiom was not stated for that case. Instead, we have the decomposition axiom D1 that separates a mixed gamble into the joint receipt of two not-mixed ones.

4.2. The main theorem

Theorem 5. Suppose $\mathscr{P} = \langle \mathscr{E}, \mathscr{C}, e, \mathscr{g}, \gtrsim, \oplus \rangle$ satisfies axioms J1-J7, S1-S2, R1-R3, and D1. Let U be the order-preserving function from $\mathscr{g}$ onto the reals given by theorem 2. If the binary gambles have unit representations in terms of U, then for each $E \in \mathscr{E}$, there exist functions $S^i(\cdot \mid E)$, $i = +, -$, from $\mathscr{E}_E$ into $[0, 1]$ such that for $g, h \in \mathscr{g}$:

(i) if g is a first-order gamble on event E and g^i is defined as in equation (6), then

$$U(g) = A(+)U(g^+)S^+[E(+) \mid E] + B(-)U(g^-)S^-[E(-) \mid E]; \tag{7}$$

(ii) if g^i is a first-order gamble on event E all of whose consequences are of sign i, where $i = +, -$, E_k is the event to which is associated the outcome nearest to e, g' and g'' are defined as in theorem 4/ii, and g'^i and g''^i are the nonnull parts defined over $E - E_k$, then

$$\begin{aligned} U(g^i) = {} & A(i)U(g'^i)S^i(E - E_k \mid E) + U[g^i(E_k)] + \\ & [C(i)/B(i)]U(g'^i)S^i(E - E_k \mid E)U[g^i(E_k)] \end{aligned} \tag{8a}$$

$$\begin{aligned} = {} & U[g^i(E_k)] + B(i)U(g''^i)S^i(E - E_k \mid E) + \\ & [C(i)/A(i)]U(g''^i)S^i(E - E_k \mid E)U[g^i(E_k)]. \end{aligned} \tag{8b}$$

The function U is unique up to a similarity transformation, and the S^i are unique. Axioms R1-R3 and D1 are consequences of this representation.

It is easy to verify from equations (8a) and (8b) that

$$A(i)U(g'^i) = B(i)U(g''^i).$$

As stated, this theorem may not appear to yield a representation for each first-order gamble, but in fact one is implied, as stated in the corollary below for the special case of $\oplus$ being additive, i.e., in its U representation $A(i) = B(i) = 1$ and $C(i) = 0$. (We omit the cumbersome expression for the general case.) The idea is simple enough. One uses part (i) to partition the problem into positive and negative components, and then one uses part (ii) recursively to simplify each of the two halves. It is clear from this recursion why the rank dependence arises. One removes the consequences sequentially, beginning with the ones nearest in preference to e, and works toward the extremes. The weights that are generated depend upon this order in an obvious way. To state the result reasonably compactly entails some notation. Let g be a first-order gamble defined on the partition $\{E_1, \ldots, E_{m-1}, E_m, E_{m+1}, \ldots, E_n\}$, where the labeling has been selected so that $g(E_j) > g(E_{j+1}), j = 1, \ldots, n-1$, and $g(E_m) = e$. Define

$$G^+(r) = S^+\left(\bigcup_{k=1}^{r-1} E_k \mid \bigcup_{k=1}^{r} E_k\right), r = 2, \ldots, m-1$$

$$F^+(j, m-1) = \begin{cases} 0, & j = 0 \\ \prod_{r=j+1}^{m-1} G^+(r), & j = 1, 2, \ldots m-2 \\ 1, & j = m-1 \end{cases}$$

$$G^-(r) = S^-\left(\bigcup_{k=r+1}^{n} E_k \mid \bigcup_{k=r}^{n} E_k\right), r = m+1, \ldots, n-1$$

$$F^-(j,m+1) = \begin{cases} 0, & j = n+1 \\ \prod_{r=m+1}^{j-1} G^-(r), & j = m+2, \ldots, n, \\ 1, & j = m+1. \end{cases}$$

Corollary to theorem 5. Suppose in the U representation of $\oplus$, $A(i) = B(i) = 1$ and $C(i) = 0$. Let g be a first-order gamble with the events labeled as above; then

$$U(g) = \left[\sum_{j=1}^{m-1} U[g(E_j)]W^+(E_j)\right] S^+[E(+) \mid E] + \left[\sum_{j=m+1}^{n} U[g(E_j)]W^-(E_j)\right] S^-[E(-) \mid E], \tag{9}$$

where

$$W^+(E_j) = F^+(j,m-1) - F^+(j-1,m-1),$$
$$W^-(E_j) = F^-(j,m+1) - F^-(j+1,m+1).$$

It is almost immediate that the sum of these weights over $E(+)$ and $E(-)$, respectively, is 1 and each is positive.

The above representation of rank dependence is somewhat similar to those of previous theories, including Gilboa (1987), Luce (1988), Quiggin (1982), Schmeidler (1989), Wakker (1989), and Yaari (1987), but it differs in one interesting way. The similarity lies in the fact that the W^i weights are expressed as a difference of two terms that are identical except for the fact that one includes the event in the question and the other does not (see p. 109 of Wakker (1989), definition VI.2.3 and equation VI.2.7, who notes its relation to the Choquet (1953–1954) integral). What differs is that the terms F^i exhibit an internal structure, which is not true of the previous theories for uncertainty, but their structure is not as some increasing function of sums of the probabilities of the events, as is true for theories of the risky case. Rather the structure is as products of conditional weights. As we discuss in the next section, whereas the formulation of Luce (1988) forces the monotonicity of general gambles, the present theory does not automatically do so. When we explore the conditions that lead to monotonicity, the present representation becomes very similar to the one found for risk.

4.3. General monotonicity

Recall that axiom R1 invokes monotonicity only for binary gambles. One must, therefore, consider the conditions under which monotonicity holds in general. Luce (1988) worked out the conditions that any rank-dependent weights must satisfy for monotonicity to hold in the case of the corollary to theorem 5. They are as follows. Consider any two permutations that differ only in the interchange of two adjacent events in the rank order.

Then all weights for events other than these two must be pairwise equal. A similar result for the general case of theorem 5 can be expressed in terms of properties of the weights.

Theorem 6. Assuming theorem 5 and treating gambles as functions over partitions, the representation is strictly monotonic increasing if and only if there exist functions $S^i, i = +, -$, from $\mathfrak{S}$ into [0, 1] such that for all $C, D \in \mathfrak{S}$ with $C \subseteq D$,

$$S^i(C \mid D) = S^i(C)/S^i(D). \tag{10}$$

Corollary. In the strictly monotonic increasing case with $A(i) = B(i) = 1$ and $C(i) = 0$, the weights in the corollary to theorem 5 take the following form:

$$W^+(E_j) = \left[S^+\left(\bigcup_{k=1}^{j} E_k\right) - S^+\left(\bigcup_{k=1}^{j-1} E_k\right)\right] / S^+\left(\bigcup_{k=1}^{m-1} E_k\right)$$

$$W^-(E_j) = \left[S^-\left(\bigcup_{k=j}^{n} E_k\right) - S^-\left(\bigcup_{k=j+1}^{n} E_k\right)\right] / S^-\left(\bigcup_{k=m+1}^{n} E_k\right).$$

This form is very similar to the one usually arrived at. Specifically, if S^i is a strictly increasing function from [0, 1] onto [0, 1] applied to a probability measure over the events, it becomes identical to the one typically derived.

Observe that if S^i is additive over disjoint events, then the W weights are of the following form:

$$W^i(E_j) = S^i(E_j)/S^i[E(i)].$$

4.4. *Estimation of parameters*

It is obvious from the definitions that if the S's are known, then the W's are determined. It is slightly less obvious that if the W's are known, then the S's are uniquely determined, but an easy calculation shows that

$$G^+(m-j) = \begin{cases} 1 - W^+(E_{m-1}), & \text{if } j = 1, \\ 1 - W^+(E_{m-j}) / \prod_{r=m-j+1}^{m-1} G(r), & \text{if } j > 1. \end{cases}$$

A similar formula holds for the negative cases. As shown by Manders in Luce (1988), in the fully monotonic case there are $2^{m-1} - 2$ independent weights for the class of gambles based on a fixed partition into $m - 1$ events. The corollary to theorem 5 easily reveals why this is so. The S^+ values must be given for all 2^{m-1} unions that can be formed, save for the null set and the set being partitioned.

Consider any two events D and E, with $D \subseteq E$, and a consequence of $c \in \mathscr{C}$ with $c > e$. Let g be the gamble with $g(D) = c$ and $g(E - D) = e$. Then by part (i) of theorem 4,

$$U(g) = U(c)S^{+}(D \mid E).$$

Similarly, for $c < e$,

$$U(g) = U(c)S^{-}(D \mid E).$$

Consider, then, the set of binary gambles with just two consequences, one of which is no change from the status quo, that are defined on partitions $\{D, E - D\}$ of a fixed event E. As has just been shown, this family must form an additive conjoint structure. In the general monotone case, the representation simplifies to

$$U(g) = U(c)S^{i}(D)/S^{i}(E), i = + \text{ if } c > e \text{ and } i = - \text{ if } c < e,$$

which is additive on three factors. We know how to estimate the functions for that situation - there is a well-developed theory (see Krantz et al., 1971, chapters 6 and 9; Wakker, 1989) as well as algorithms that have been incorporated into computer packages. One early study of such gambles from a conjoint measurement perspective is Tversky (1967). He concluded that either the multiplicative form fails to hold or the S parameters are not additive over disjoint subevents, as they would be if they were probabilities. The latter finding is completely consistent with the representation being rank dependent. Once both the utility function and the S weights have been estimated, the theory makes numerous predictions that can be tested.

5. Conclusions

The major conclusion of this article is that it is possible to give a relatively simple axiomatization of a theory of preferences among uncertain alternatives that makes an explicit distinction between gains and losses and that generalizes both prospect theory and rank-dependent theory. The axiomatization rests critically on having, in addition to the structure of gambles, an operation of joint receipt, which allows explicit editing of gambles. That editing process, which proceeds inductively beginning with the consequence closest to the status quo, is the source of the rank dependence in the domains both of all gains and of all losses. Moreover, a specific axiomatization of joint receipt yields a representation that is interestingly unusual. The major source of nonrationality in the theory—which is bound to exist if it is not to reduce to SEU or to pure rank dependence—is the assumption that a gamble of gains and losses is treated as the joint receipt of two gambles, one being the gains pitted against no change from the status quo and the other of the losses pitted against no change from the status quo. This property, which is implicit in prospect theory, seems somewhat plausible, but is surely in need of empirical study.

Appendix

Proofs

All proofs for the negative quadrant are omitted as being similar to those for the positive quadrant.

Proof of Theorem 1. Throughout we suppress both the j (sign) superscripts and the notation for E on $\otimes$, G, and S_i^j. Moreover, the assumption that $\oplus$ is accumulative implies $x \oplus y \geq y$, which is used implicitly in invoking equation (2).

(i) + (ii) $\Rightarrow$ (iii):

$$\begin{aligned} U[(x \otimes 0) \oplus y] &= AU(x \otimes 0) + BU(y) + CU(x \otimes 0)U(y) \\ &= [A + CU(y)][S_> U(x)] + BU(y) \\ &= AS_> U(x) + BU(y) + CS_> U(x)U(y). \end{aligned}$$

Since $x \oplus y \geq y$ and $U(0) = 0$,

$$\begin{aligned} U[(x \oplus y) \otimes (0 \oplus y)] &= S_>)U(x \oplus y) + (1 - S_>)U(0 \oplus y) \\ &= [AU(x) + BU(y) + CU(x)U(y)]S_> + (1 - S_>)BU(y) \\ &= AS_> U(x) + BU(y) + CS_> U(x)U(y). \end{aligned}$$

So they are equal, and applying U^{-1} yields the left half of (iii). The proof of the right half is similar.

(ii) + (iii) $\Rightarrow$ (i):[7]

Denote $\oplus$ as a function F, apply (ii)(a) to the left side of (iii), and invoke equation (3):

$$\begin{aligned} F(x \otimes 0, y) &= F(U^{-1}[S_> U(x) + (1 - S_>)U(0)], y) \\ &= U^{-1}[S_> UF(x,y) + (1 - S_>)F(0,y)]. \end{aligned}$$

Set $X = U(x)$, $Y = U(y)$, and $H(X, Y) = UF[U^{-1}(X), U^{-1}(Y)]$, and use $U(0) = 0$ to get

$$H(S_> X, Y) = S_> H(X, Y) + (1 - S_>)H(0, Y).$$

Rewriting,

$$H(S_> X, Y) - H(0, Y) = S_> [H(X, Y) - H(0, Y)].$$

Since, by (ii)(b), $S_>$ can be any number in the interval (0, 1), it is well known (Aczél, 1966) that the strictly increasing, onto solutions are

$$H(X, Y) = X\alpha(Y) + H(0, Y), \tag{11}$$

with $\alpha > 0$. A similar argument applied to the right side of (iii) yields

$$H(X, Y) = \beta(X)Y + H(X, 0), \tag{12}$$

with $\beta > 0$. By (ii)(c),

$$H(0, 0) = UF[U^{-1}(0), U^{-1}(0)] = U(0 \oplus 0) = U(0) = 0.$$

Substituting this into equations (11) and (12),

$$H(X, 0) = AX \text{ and } H(0, Y) = BY,$$

where $A = \alpha(0) > 0$ and $B = \beta(0) > 0$. Thus,

$$H(X, Y) = X\alpha(Y) + BY = \beta(X)Y + AX.$$

Rewriting,

$$[\alpha(U(y)) - A]/U(y) = [\beta(U(x)) - B]/U(x).$$

So, either $\alpha = A$ and $\beta = B$, which is case (i)(a) with $C = 0$, or for all y

$$[\alpha(U(y)) - A]/U(y) = C.$$

This substituted in equation (11) yields (i)(a) with $C \neq 0$.
(i)(b) is obvious from (ii)(a), and (i)(c) from (ii)(b).

(i) + (iii) $\Rightarrow$ (ii)
Suppose $x > y$. Let v solve $y = 0 \oplus v$ and u solve $x = u \oplus v$. The former exists because U is onto Re^+ and $y \geq 0$, the latter because $\oplus$ is accumulative and so $x > y = 0 \oplus v \geq v$. Observe that, by equation (2), $U(y) = BU(v)$ and $U(x) = AU(u) + BU(v) + CU(u)U(v)$, and so $U(u) = [U(x) - U(y)]/[A + (C/B)U(y)]$. Using assumption (iii), the assumption that 0 is a generalized zero, and the above substitutions,

$$\begin{aligned} U(x \otimes y) &= U[(u \oplus v) \otimes (0 \oplus v)] \\ &= U[(u \otimes 0) \oplus v] \\ &= AU(u \otimes 0) + BU(v) + CU(u \otimes 0)\, U(v) \\ &= KU(u)[A + CU(v)] + BU(v) \\ &= K[U(y) - U(x)] + U(y). \end{aligned}$$

Identifying K with $S^{>j}$ yields the upper part of equation 3. The lower part is proved similarly.
Property (ii)(b) follows from (i)(b).

Proof of Corollary to Theorem 1. The only thing that needs to be modified is the proof that (i) + (iii) ⇒ (ii). From assumption (iii) and $0 \oplus 0 = 0$,

$$(x \otimes 0) \oplus 0 = (x \oplus 0) \otimes (0 \oplus 0) = (x \oplus 0) \otimes 0.$$

Thus, by (i)(a) and the fact that the monotonicity of $\otimes$ implies the existence of a strictly increasing function F such that $U(x \otimes 0) = F[U(x)]$,

$$AF[U(x)] = AU(x \otimes 0) = U[(x \oplus 0) \otimes 0] = \mathrm{F}\,[U(x \oplus 0)] = F[AU(x)].$$

Assuming $A \neq 1$, it is well known that the general solution to this equation is

$$F(Z) = kZ \exp[p(\log Z)],$$

where p is periodic with period $\log A$.

Suppose $x > y$. As in the preceding proof, let v solve $y = 0 \oplus v$ and u solve $x = u \oplus v$. Writing $G(Z) = k\exp[p(\log Z)]$ and substituting,

$$\begin{aligned} U(x \otimes y) = U[(u \oplus v) \otimes (0 \oplus v)] &= U[(u \otimes 0) \oplus v] \\ &= AU(u \otimes 0) + BU(v) + CU(u \otimes 0)\,U(v) \\ &= AU(u)G[U(u)] + BU(v) + \\ &\quad\; CU(u)G[U(u)]U(v) \\ &= G\left[\frac{U(x)-U(y)}{A+(C/B)U(y)}\right][U(x) - U(y)] + U(y). \end{aligned}$$

To get the desired result, we must show that G is in fact constant. Supposing otherwise, we establish that the function

$$H(X, Y) = G([X - Y]/[A + (C/B)Y])(X - Y) + Y$$

is not strictly increasing in Y, which contradicts the monotonicity assumption about $\otimes$. Let $Z = (X - Y)/[A + (C/B)Y]$, which is strictly decreasing with Y because with $A > 0$, $B > 0$, $CX > 0$, and so

$$\frac{\partial Z}{\partial Y} = -\frac{A+(C/B)X}{[A+(C/B)Y]^2} < 0.$$

Solving for $Y = (X - AZ)/[1 + (C/B)Z]$ and substituting,

$$H(X, Y) = \frac{G(Z)Z[A+(C/B)X]+X-AZ}{1+(C/B)Z}.$$

Select $Z = A^n Z_0$. Then, since p is periodic with period $\log A$,

$$G(A^n Z_0) = k \exp[p(\log A^n Z_0)] = k \exp[p(\log Z_0)] = G(Z_0).$$

Thus, defining Y_n by $A^n Z_0 = (X - Y_n)/[A + (C/B)Y_n]$, we have

$$H(X,Y_n) = \frac{G(Z_0)A^n Z_0[A+(C/B)X]+X-A^{n+1}Z_0}{1+(C/B)A^n Z_0}.$$

If $A > 1$, let n become large, and if $A < 1$, let n approach $-\infty$. Then the preceding expression approaches $(B/C)G(Z_0)[A+(C/B)X] - AB/C$. Since p is periodic, if it is not constant, then we may find Z_0 and Z_1 such that $Z_0 > Z_1$ and $p(Z_0) > p(Z_1)$, and so $G(Z_0) > G(Z_1)$. But since Z is a decreasing function of Y, this means that for the corresponding Y values, the order is inverted, which violates monotonicity.

The proof for $x < y$ is similar.

If $A = 1$ and $B \neq 1$, then a similar argument using the second half of property (iii) yields the same conclusion.◇

Proof of Theorem 2. Assume that axioms J1-J7 hold. By a standard result, axiom J1 implies that there is an F: $\mathcal{g} \to Re$ such that

$$\forall g, h \in \mathcal{g}, g \gtrsim h \Leftrightarrow F(g) \geq F(h), \tag{13}$$

with $F(e) = 0$ and $F(\mathcal{g})$ an interval that includes both positive and negative numbers. Suppose for some g, $F(g) = \sup F(\mathcal{g})$. Take $e > h$ by axiom J1 and conclude from axiom J3 that $g \oplus h \sim x \oplus e$ for some $x \in \mathcal{g}$. Then axiom J2 implies $x > g$, contrary to equation (13) and the supposed maximality of g. Hence $\mathcal{g}$ has no maximal element and, similarly, no minimal element. With no loss of generality we assume henceforth that $F(\mathcal{g}) = Re$.

For each t in the set $\mathcal{g}/\sim$ of equivalence classes of $\mathcal{g}$ under $\sim$, define $F(t)$ as the common value of $F(g)$ for all $g \in t$. Since axiom J2 allows free substitution of equivalent elements in expressions involving $\oplus$ without affecting $\gtrsim$ or F, we can translate axioms J2–J7 into an equivalent real system in which $F(t)$ is the surrogate of every $g \in t$. In the new system, $e \to 0$, $\mathcal{g}^+ \to R^+$ and $\mathcal{g}^- \to R^-$, where

$$R^+ = Re^+ \cup \{0\} \text{ and } R^- = Re^- \cup \{0\}.$$

With the abuse of notation noted just before theorem 1, the translated axioms are as follows:

J2*. $\oplus$ on Re is closed, $0 \oplus 0 = 0$, and $\forall p, q, r \in Re$, if $p > q$ then $p \oplus r > q \oplus r$ and $r \oplus p > r \oplus q$.

J3*. Given any three of the four terms in any one of $p \oplus q = r \oplus s$, $(p \oplus q) \oplus r = s$, and $r \oplus (p \oplus q) = s$, the fourth term exists so that the equality holds.

J4*. With $(p, q) \in Re^2$ defined as $p \oplus q$, $\langle R^+ \times R^-, \geq \rangle$ and $\langle R^- \times R^+, \geq \rangle$ are additive conjoint structures.

J5*. $\forall a, b, c \in R^+, \forall x, y, z \in R^-$,

(i) if $a \oplus 0 = b \oplus x, 0 \oplus a = y \oplus b$ and $c \oplus 0 = a \oplus x$, then $0 \oplus c = y \oplus a$;
(ii) if $x \oplus 0 = y \oplus a, 0 \oplus x = b \oplus y$ and $z \oplus 0 = x \oplus a$, then $0 \oplus z = b \oplus x$.

J6*. $\forall a, b \in R^+, \forall x, y \in R^-$, any two of the following imply the third:

$$a \oplus 0 = 0 \oplus b, 0 \oplus x = y \oplus 0, a \oplus x = y \oplus b.$$

J7*. $\forall a, b, c, d, g \in R^+, \forall x, y, z, w, u \in R^-$,

(i) If $b \oplus 0 = a \oplus x, d \oplus 0 = c \oplus x$ and $(b \oplus g) \oplus 0 = (a \oplus g) \oplus y$, then $(d \oplus g) \oplus 0 = (c \oplus g) \oplus y$;
(ii) If $y \oplus 0 = x \oplus a, w \oplus 0 = z \oplus a$ and $(y \oplus u) \oplus 0 = (x \oplus u) \oplus b$, then $(w \oplus u) \oplus 0 = (z \oplus u) \oplus b$.

The two expressions obtained by commuting every $\oplus$ term in (i) and (ii) also hold.

We will work with axioms J2*–J7* until the last paragraph of the proof. For notational ease, a, b, c, and d with or without appurtenances, *always* denote elements in R^+, and x, y, and z *always* denote elements in R^-. We often omit $\forall$ when it clearly applies.

A few lemmas will guide our way to the representation of equation (1). We begin with the mixed domains.

Lemma 1. There exist strictly increasing ϕ and ψ from Re onto Re with $\phi(0) = \psi(0) = 0$ such that, $\forall a, b \in R^+, \forall x, y \in R^-$,

$$\begin{aligned} a \oplus x > b \oplus y &\Leftrightarrow \phi(a) + \phi(x) > \phi(b) + \phi(y) \\ x \oplus a > y \oplus b &\Leftrightarrow \psi(x) + \psi(a) > \psi(y) + \psi(b). \end{aligned} \tag{14}$$

Each of ϕ and ψ is unique up to a similarity transformation.

Proof. Axioms J4* and J2* imply strictly increasing ϕ_1 on R^+ and ϕ_2 on R^- such that

$$a \oplus x > b \oplus y \Leftrightarrow \phi_1(a) + \phi_2(x) > \phi_1(b) + \phi_2(y).$$

Fix origins by specifying $\phi_1(0) = \phi_2(0) = 0$, then let ϕ equal ϕ_1 on R^+ and ϕ_2 on R^- to obtain equation (14). It follows without difficulty from axioms J2* and J3* that $\phi(Re)$ and $\psi(Re)$ are intervals that in fact equal Re. The uniqueness result with fixed origins is well known. $\diamond$

Fix ϕ and ψ henceforth. When the two $\Leftrightarrow$ expressions of lemma 1 are used in each part of axiom J5*, its two parts yield

$$\phi(a) = [\phi(b) + \phi(c)]/2 \Leftrightarrow \psi(a) = [\psi(b) + \psi(c)]/2$$
$$\phi(x) = [\phi(y) + \phi(z)]/2 \Leftrightarrow \psi(x) = [\psi(y) + \psi(z)]/2$$

respectively. It follows from these and lemma 1 that there are positive constants λ and μ such that

$$\forall a \in R^+, \psi(a) = \lambda\phi(a)$$
$$\forall x \in R^-, \psi(x) = \mu\phi(x).$$

Hence,

$$x \oplus a > y \oplus b \Leftrightarrow \mu\phi(x) + \lambda\phi(a) > \mu\phi(y) + \lambda\phi(b). \tag{15}$$

Define U on $\{a \oplus x\} \cup \{x \oplus a\}$ by

$$U(a \oplus x) = \phi(a) + \phi(x)$$
$$U(x \oplus a) = \tau[\mu\phi(x) + \lambda\phi(a)],$$

where $\tau > 0$ is a constant to be determined so that $U(a \oplus x) = U(y \oplus b)$ whenever $a \oplus x = y \oplus b$.

Lemma 2. There is a unique $\tau > 0$ such that

$$\phi(a) + \phi(x) = \tau[\mu\phi(y) + \lambda\phi(b)] \text{ whenever } a \oplus x = y \oplus b.$$

Proof. Fix positive a_0, b_0 and negative x_0, y_0 so that $a_0 \oplus 0 = 0 \oplus b_0$ and $a_0 \oplus x_0 = 0 = y_0 \oplus b_0$, as assured by axiom J3* and monotonicity. Axiom J6* implies $0 \oplus x_0 = y_0 \oplus 0$. Define τ by

$$\phi(a_0) = \tau\lambda\phi(b_0)$$

so that $U(a_0 \oplus 0) = U(0 \oplus b_0)$. By equation (14), $\phi(a_0) + \phi(x_0) = 0$; by equation (15), $\mu\phi(y_0) + \lambda\phi(b_0) = 0$; hence

$$\phi(x_0) = \tau\mu\phi(y_0),$$

so that $U(0 \oplus x_0) = U(y_0 \oplus 0)$ goes along with $0 \oplus x_0 = y_0 \oplus 0$. Moreover, whenever $a \oplus x = 0 = y \oplus b$, we have

$$U(a \oplus x) = \phi(a) + \phi(x) = 0 = \tau[\mu\phi(y) + \lambda\phi(b)] = U(y \oplus b).$$

The other two general cases we must verify are

$$a \oplus x = y \oplus b > 0 \Rightarrow U(a \oplus x) = U(y \oplus b),$$
$$a \oplus x = y \oplus b < 0 \Rightarrow U(a \oplus x) = U(y \oplus b).$$

We prove the first of these.

Suppose $a \oplus x = y \oplus b > 0$. Let c and d satisfy axiom J3*:

$$c \oplus 0 = a \oplus x \text{ and } 0 \oplus d = y \oplus b.$$

By equations (14) and (15), $\phi(a) + \phi(x) = \phi(c)$ and $\lambda\phi(d) = \mu\phi(y) + \lambda\phi(b)$, so to verify $U(a \oplus x) = U(y \oplus b)$ it suffices to show that $\phi(c) = \tau\lambda\phi(d)$. This reduces the first case of the preceding paragraph to

Claim 1. $\forall c, d > 0, c \oplus 0 = 0 \oplus d \Rightarrow \phi(c) = \tau\lambda\phi(d)$.

We verify this by construction, following a process in Fishburn (unpublished) that uses uniform sequences and bisections.

Given a_0, b_0, x_0, y_0 as above, define a_i, b_i for $i = 1, 2, \ldots$, recursively by

$$a_i \oplus x_0 = a_{i-1} \oplus 0 = 0 \oplus b_{i-1} = y_0 \oplus b_i$$

with $a_i \oplus 0 = 0 \oplus b_i$ by axiom J6*. For $i = 1$ we have

$$\phi(a_1) + \phi(x_0) = \phi(a_0) = \tau\lambda\phi(b_0) = \tau[\mu\phi(y_0) + \lambda\phi(b_1)], \phi(x_0) = \tau\mu\phi(y_0),$$

where the first and third equalities are from equations (14) and (15), and the others were noted previously. It follows that

$$\phi(a_1) = \tau\lambda\phi(b_1) = 2\phi(a_0).$$

Continuation yields

$$\forall i \geq 0, a_i \oplus 0 = 0 \oplus b_i \text{ and } \phi(a_i) = \tau\lambda\phi(b_i) = (i + 1)\phi(a_0).$$

We refer to this as *the uniform sequence based on a_0 and b_0*.

For bisection that begins from a_0 and b_0, define a and x by

$$a_0 \oplus x = a \oplus 0 \text{ and } a \oplus x = 0.$$

Also define b and y by

$$a \oplus 0 = 0 \oplus b \text{ and } 0 \oplus b = y \oplus b_0.$$

We have $0 \oplus x = y \oplus 0$ and $y \oplus b = 0$ by axiom J6*. Therefore, by equations (14) and (15),

$$\begin{aligned} \phi(a_0) + \phi(x) &= \phi(a) = -\phi(x) \\ \mu\phi(y) + \lambda\phi(b_0) &= \lambda\phi(b) = -\mu\phi(y). \end{aligned}$$

These imply that $\phi(a) = \phi(a_0)/2$ and $\phi(b) = \phi(b_0)/2$. Hence, since $\phi(a_0) = \tau\lambda\phi(b_0)$,

$$\phi(a) = \tau\lambda\phi(b) \text{ and } a \oplus 0 = 0 \oplus b.$$

(They also lead to $\phi(x) = \tau\mu\phi(y)$ and $0 \oplus x = y \oplus 0$.) Continued bisection produces sequences $a = a_{-1}, a_{-2}, \ldots$ and $b = b_{-1}, b_{-2}, \ldots$ such that

$$\phi(a_{-i}) = \tau\lambda\phi(b_{-i}) = \phi(a_0)/2^i \text{ and } a_{-i} \oplus 0 = 0 \oplus b_{-i}.$$

The final step in the construction builds the uniform sequence based on a_{-i} and b_{-i}, $i \geq 1$, in the manner done earlier for a_0 and b_0. Omitting routine details (cf. Fishburn, unpublished), it follows that claim 1 holds for a set of $c \in R^+$ whose $\phi(c)$ values are dense in R^+. Monotonicity and the onto property for ϕ complete the proof of the claim.◇

We have arrived at the following point: There exist strictly increasing maps ϕ and U from *Re* onto *Re* and positive constants λ, μ, and τ such that $\phi(0) = U(0) = 0$,

$$U(a \oplus x) = \phi(a) + \phi(x) \qquad \text{on } \{a \oplus x\} \tag{16a}$$

$$U(x \oplus a) = \tau\mu\phi(x) + \tau\lambda\phi(a) \qquad \text{on } \{x \oplus a\}, \tag{16b}$$

and, $\forall p, q \in \{a \oplus x\} \cup \{x \oplus a\}$,

$$p > q \Leftrightarrow U(p) > U(q) \Leftrightarrow \phi(p) > \phi(q).$$

Next we show that, on each of R^+ and R^-, ϕ is a similarity transformation of U.

Lemma 3. There are positive α and β such that

$$\forall a \in R^+, \phi(a) = \alpha U(a),$$
$$\forall x \in R^-, \phi(x) = \beta U(x).$$

Proof. To establish the R^+ result, we show first that for all $a > b > 0, c > d > 0$, and $g \in R^+$,

$$\frac{\phi(a \oplus g) - \phi(b \oplus g)}{\phi(c \oplus g) - \phi(d \oplus g)} = \frac{\phi(a) - \phi(b)}{\phi(c) - \phi(d)}. \tag{17}$$

Suppose $\phi(a) - \phi(b) = \phi(c) - \phi(d) > 0$. By axiom J3* and equation (14), there is an $x \in R^-$ for which $b \oplus 0 = a \oplus x$ and $d \oplus 0 = c \oplus x$. Given any $g \in R^+$, $a \oplus g > b \oplus g$ by axiom J2*, so axioms J2* and J3* yield a $y \in R^-$ such that $(b \oplus g) \oplus 0 = (a \oplus g) \oplus y$. Then $(d \oplus g) \oplus 0 = (c \oplus g) \oplus y$ by part (i) of axiom J7*. Hence, by equation (14), $\phi(a \oplus g) - \phi(b \oplus g) = -\phi(y) = \phi(c \oplus g) - \phi(d \oplus g)$, so equation (17) holds when its right side equals 1.

Suppose the right side of equation (17) is a rational number, say r/s in the smallest positive integers. Partition $\phi(a) - \phi(b)$ into r equal and contiguous segments; do likewise for $\phi(c) - \phi(d)$ into s segments. Then every segment has the same length, and it follows from similar partitions for the left side and the result of the preceding paragraph that $[\phi(a \oplus g) - \phi(b \oplus g)]/[\phi(c \oplus g) - \phi(d \oplus g)] = r/s$. The strictly increasing and onto properties for ϕ then imply that equation (17) holds whenever $a > b > 0, c > d > 0$, and $g \in R^+$.

Set $g = 0$ and $b = \mathrm{d}$ in equation (17) to obtain

$$\frac{\phi(a \oplus 0) - \phi(b \oplus 0)}{\phi(c \oplus 0) - \phi(b \oplus 0)} = \frac{\phi(a) - \phi(b)}{\phi(c) - \phi(b)}$$

for $\min\{a, c\} > b > 0$. Using the form $(p \oplus q) \oplus r = s$ of axiom J3*, let $q = s = 0$, take $r < 0$, and conclude from axiom J3* that $(b \oplus 0) \oplus r = 0 = 0 \oplus 0$ for some $b > 0$. By equation (14), $\phi(b \oplus 0) = -\phi(r)$. Let a sequence of $\phi(r)$ values approach 0 from below. Then the corresponding sequence of $\phi(b \oplus 0)$ values approaches 0 from above. Using axiom J2*, it follows easily that the b sequence, say $b_1 > b_2 > \ldots$, approaches 0, and so $\phi(b_i) \to 0$ also. We conclude from the preceding ϕ ratio equation that

$$\forall a, c > 0, \frac{\phi(a \oplus 0)}{\phi(c \oplus 0)} = \frac{\phi(a)}{\phi(c)}.$$

By equation (16a), $U(a \oplus 0) = \phi(a)$ and $U(c \oplus 0) = \phi(c)$, so

$$\frac{\phi(a \oplus 0)}{U(a \oplus 0)} = \frac{\phi(c \oplus 0)}{U(c \oplus 0)}$$

or when $b = a \oplus 0$ and $d = c \oplus 0$, $\phi(b)/U(b) = \phi(d)/U(d)$. Since it is easily seen that for every $b > 0$ there is an $a > 0$ such that $b = a \oplus 0$, it follows that $\phi(b)/U(b)$ is the same for all $b > 0$. We define the common ratio to be α, yielding the first conclusion of lemma 3.◇

Given lemma 3, let $A(+) = \alpha, B(-) = \beta, A(-) = \tau\mu\beta$, and $B(+) = \tau\lambda\alpha$. Then, by equation (16),

$$U(a \oplus x) = A(+)U(a) + B(-)U(x) \quad \text{on } \{a \oplus x\} \tag{18a}$$

$$U(x \oplus a) = A(-)U(x) + B(+)U(a) \quad \text{on } \{x \oplus a\}. \tag{18b}$$

It remains to establish the real equivalents of the first and fourth lines of equation (1b), i.e.,

$$U(a \oplus b) = A(+)U(a) + B(+)U(b) + C(+)U(a)U(b)$$
$$U(x \oplus y) = A(-)U(x) + B(-)U(y) + C(-)U(x)U(y)$$

With $C(+) \geq 0$ and $C(-) \leq 0$. We prove the first of these, beginning with

Lemma 4. $\forall a, b, c, d, f \in R^+$ with $a > b$ and $c > d$

$$\frac{U(a \oplus f) - U(b \oplus f)}{U(c \oplus f) - U(d \oplus f)} = \frac{U(a) - U(b)}{U(c) - U(d)} = \frac{U(f \oplus a) - U(f \oplus b)}{U(f \oplus c) - U(f \oplus d)}.$$

Proof. The first equality follows from equations (16a) and (17) and from lemma 3. The second follows similarly using the commuted version of equation (17), which is obtained from the commuted version of part (i) of axiom J7*.◇

Now fix $a_0 > b_0 > 0$, define r_1 and r_2 on R^+ by

$$r_1(f) = \frac{U(a_0 \oplus f) - U(b_0 \oplus f)}{U(a_0) - U(b_0)} > 0$$

$$r_2(f) = \frac{U(f \oplus a_0) - U(f \oplus b_0)}{U(a_0) - U(b_0)} > 0$$

and use equation (18) and lemma 4 with $d = 0$ to get

$$U(c \oplus f) = r_1(f)U(c) + B(+)U(f)$$

$$U(c \oplus f) = r_2(c)U(f) + A(+)U(c),$$

where c and f have been interchanged for the latter equation. Equating the right sides of these two $U(c \oplus f)$ equations gives

$$\frac{r_1(f) - A(+)}{U(f)} = \frac{r_2(c) - B(+)}{U(c)}.$$

Define this common ratio for all $c > 0$ as $C(+)$. Then, when $C(+)U(f) + A(+)$ is substituted for $r_1(f)$, we get

$$U(c \oplus f) = A(+)U(c) + B(+)U(f) + C(+)U(c)U(f),$$

which is the desired result. If $C(+)$ were negative, suitably large values of $U(c)$ and $U(f)$ would give $U(c \oplus f) < 0$, a contradiction since U is strictly increasing and $c \oplus f > 0$ by axiom J2*. Hence $C(+) \geq 0$.

The similar proof for $U(x \oplus y)$ requires $C(-) \leq 0$.

This completes the sufficiency proof of our real-domain translation for the representation of equation (1), along with part (i) of the corollary of theorem 2. Reversal into the original domain for theorem 2 as stated is straightforward. We omit the simple proofs of parts (ii) and (iii) of the corollary.◇

Proof of Theorem 3. By the corollary to theorem 2, $A(j) = B(j) = 1$. Observe that for $C = 0$, U is additive over $\oplus$, and Narens (1981) established that the isomorphism between two unit structures must be a power function. So the form for U, equation (4), follows immediately. For $C \neq 0$, let $V = \log(CU + 1)$ and observe

$$\begin{aligned} V(x \oplus y) &= \log[CU(x \oplus y) + 1] \\ &= \log[CU(x) + CU(y) + C^2U(x)U(y) + 1] \\ &= \log[(CU(x) + 1)(CU(y) + 1)] \\ &= V(x) + V(y). \end{aligned}$$

Thus, as in the $C = 0$ case, V is a power function, yielding for U the form of equation (5).◇

Proof of Theorem 4. (i) Suppose $b > c \gtrsim e$, $c, b \in \mathscr{C}$. Since U is onto, there exists by axiom S1 $c', c'' \in \mathscr{C}$ such that $U(c) = B(+)U(c') = A(+)U(c'')$, in which case by theorem 2, $c \sim e \oplus c' \sim c'' \oplus e$. Let d' solve

$$\begin{aligned} U(d') &= [U(b) - B(+)U(c')]/[A(+) + C(+)U(c')] \\ &> [U(c) - B(+)U(c')]/[A(+) + C(+)U(c')] = 0. \end{aligned}$$

Thus,

$$U(b) = A(+)U(d') + B(+)U(c') + C(+)U(d')U(c') = U(d' \oplus c'),$$

whence by theorem 2, $b \sim d' \oplus c'$. The proof for c'' is similar.

(ii) g' and g'' exist by related applications of part (i), and the conclusion is an immediate consequence of axiom R3.◇

Proof of Theorem 5. Let U be the function established in theorem 2. Since the binary gambles are assumed to have unit representations in terms of U, theorem 1 implies that there exist constants $S^i(E \mid E \cup E')$ such that for $g(i) \in \mathscr{G}^i$,

$$U[g(i), E; e, E'] = U[g(i)]S^i(E \mid E \cup E'). \tag{19}$$

Part (i) now follows immediately from equation (19), the additivity of U for consequences or gambles of opposite sign (equation (1)), and axiom D1.

(ii) By theorems 2 and 4,

$$U(g^i) = A(i)U(g^{i\prime}) + B(i)U(c') + C(i)U(g^{i\prime})U(c'). \tag{20}$$

By definition, $c \sim e \oplus c'$, so by theorem 2, $U(c) = B(i)U(c')$. By axiom R2, $g^{i\prime}$ is indifferent to the binary gamble of its nonnull part g'^i over the event $E - E_k$ and e over E_k. So by theorem 1, $U(g^{i\prime}) = U(g'^i)S^i(E - E_k \mid E)$. Substituting these two expressions into equation (20) yields equation (8a).

The proof of equation (8b) is similar using subtraction from the left.

The uniqueness of U up to a similarity follows from theorem 2 and the absolute uniqueness of the weights from theorem 1.

It is trivial to verify that axioms R1–R3 and D1 follows from the representation.◇

Proof of Corollary to Theorem 5. By part (i), it is sufficient to derive the representation of positive and negative gambles separately. The proof for the positive case, which we show, proceeds by induction. Using axiom R3 and part (ii),

$$\begin{aligned} U(g^+) &= U(g')G^+(m-1) + U[g^+(E_{m-1})] \\ &= \sum_{j=1}^{m-2} U[g(E_j) \ominus g(E_{m-1})][F^+(j, m-2) - \\ &\qquad F^+(j-1, m-2)]G^+(m-1) + U[g^+(E_{m-1})] \\ &= \sum_{j=1}^{m-2} U[g(E_j)][F^+(j, m-1) - F^+(j-1, m-1)] + \\ &\qquad U[g^+(E_{m-1})]\left[1 - \sum_{j=1}^{m-2} F^+(j, m-1) + \sum_{j=1}^{m-2} F^+(j-1, m-1)\right] \\ &= \sum_{j=1}^{m-1} U[g(E_j)]W^+(E_j).\diamond \end{aligned}$$

Proof of Theorem 6. Given the bilinear forms of theorem 5, the only possibility for a failure of monotonicity is either in increasing a negative consequence through e to a positive one or, in the single signed domain, to an increase in one consequence that alters the rank ordering. The former case is, according to equation (7), not a problem since the change takes a negative term to a positive one. In the latter cases, we work out the details only for the positive case and number the events according to the ranking as in the corollary of theorem 5 from 1 to $m-1$. Suppose that i and $i+1$ are adjacent events whose order will be reversed by the change in the consequence. Observe that the induction up to event $i+1$ is independent of the change in rank between these two events, so with no loss of generality we may assume gambles g and h based on $E = \bigcup_1^{i+1} E_j$, which differ only in the rank order of the consequences on these two adjacent events. Let $C = E - E_i - E_{i+1}$, $D = E - E_{i+1}$, and $D' = E - E_i$. So, by applying equation (8a) twice and suppressing the $+$ notation, we have for g

$$\begin{aligned} U(g) = [&AU(g'')\, S(C \mid D) + U[g(E_i)] + (C/B)U(g'')S(C \mid D)] \times \\ &[A + (C/B)U(g(E_{i+1}))]S(D \mid E) + U[g(E_{i+1})]. \end{aligned}$$

The equation for $U(h)$ is similar, but with D' replacing D, the roles of E_i and E_{i+1} reversed, and h replacing g. Observe that $g'' = h''$ and at the point of inversion $g(E_i) = g(E_{i+1}) = h(E_i) = h(E_{i+1})$. So equating the two expressions at that point of inversion and taking these facts into account yields

$$S(C \mid D)S(D \mid E) = S(C \mid D')S(D' \mid E), \tag{21}$$

where $C \subseteq D \subseteq E$ and $C \subseteq D' \subseteq E$. Choose E to be the universal event of $\mathcal{E}$, and set $S(C) = S(C \mid E)$. Since the identity automorphism is represented as 1, $S(E \mid E) = 1$, and so letting $D' = E$, we obtain from equation (21) the condition that

$$S^i(C \mid D) = S^i(C)/S^i(D), \tag{22}$$

where the sign notation has been reintroduced.

Clearly, the condition of equation (22) implies equation (21), which in turn is sufficient for monotonicity, and so the assertion is proved.◇

Notes

1. To the best of our knowledge, the several nonadditive weight theories in the literature exhibit interval scale representations. In particular, the rank-dependent ones in the economic literature do so.
2. They called it *dual bilinear*, but subsequently Luce adopted the terminology that has arisen in economics for theories of weighted utility where the weights depend both on the events underlying the consequences and on the preference ranking of the consequence in question among all the consequences of that particular gamble.
3. $E_j \in \mathcal{E}_E, E_j \subseteq E, j = 1, \ldots, n$; for $j \neq k, E_j \cap E_k = \emptyset$; and $\cup E_j = E$.
4. The structure is said to be *order dense* iff for every $g, h \in \mathcal{G}$ with $g > h$, there exists $f \in \mathcal{G}$ such that $g > f > h$. The structure is said to be *Dedekind complete* iff every bounded, nonempty subset of $\mathcal{G}$ has a least upper bound in $\mathcal{G}$.
5. The motivation for such an assumption and a number of its properties can be found in Cohen and Narens (1979) and Luce and Narens (1985).
6. This is equivalent to $A(j) = B(j) = 1$ (see the corollary to theorem 2).
7. Discussions with J. Aczél have been useful concerning this portion of the proof.

References

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

Choquet, G. (1953-1954). "Theory of Capacities," *Annales de l'Institut Fourier* 5, 131-295.

Cohen, M., and L. Narens. (1979). "Fundamental Unit Structures: A Theory of Ratio Scalability," *Journal of Mathematical Psychology* 20, 193-232.

Edwards, W. (1962). "Subjective Probabilities Inferred from Decisions," *Psychological Review* 69, 109-135.

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

Fishburn, P. C. (1988). *Nonlinear Preference and Utility Theory.* Baltimore, MD: Johns Hopkins Press.

Fishburn, P. C. (unpublished). "Continuous Nontransitive Additive Conjoint Measurement," Preprint, AT&T Bell Laboratories.

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

Kahneman, D., and A. Tversky. (1979). "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica* 47, 263-291.

Keeney, R. L., and H. Raiffa. (1976). *Decisions with Multiple Objectives: Preferences and Value Tradeoffs.* New York: Wiley.

Keller, L. R. (1985). "The Effects of Problem Representation on the Sure-Thing and Substitution Principles," *Management Science* 31, 738-751.

Krantz, D. H., R. D. Luce, P. Suppes, and A. Tversky. (1971). *Foundations of Measurement*, vol. I. New York: Academic Press.

Luce, R. D. (1988). "Rank-Dependent, Subjective Expected-Utility Representations," *Journal of Risk and Uncertainty* 1, 305-332.

Luce, R. D. (In press (a)). "Rank- and Sign-Dependent Linear Utility Models for Binary Gambles," *Journal of Economic Theory.*

Luce, R. D. (In press (b)). "Where Does Subjective Expected Utility Fail Descriptively?" *Journal of Risk and Uncertainty.*

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

Narens, L. (1981). "A General Theory of Ratio Scalability with Remarks About the Measurement-Theoretic Concept of Meaningfulness," *Theory and Decision* 13, 1-70.

Pfanzagl, J. (1959). "A General Theory of Measurement—Applications to Utility," *Naval Research Logistics Quarterly* 6, 283-294.

Quiggin, J. (1982). "A Theory of Anticipated Utility," *Journal of Economic Behavior and Organization* 3, 324-343.

Schmeidler, D. (1989). "Subjective Probability and Expected Utility Without Additivity." *Econometrica* 57, 571-587.

Slovic, P. (1967). "The Relative Influence of Probabilities and Payoffs upon Perceived Risk of a Gamble," *Psychonomic Science* 9, 223-224.

Slovic, P., and S. Lichtenstein. (1968). "Importance of Variance Preferences in Gambling Decisions," *Journal of Experimental Psychology* 78, 646-654.

Tversky, A. (1967). "Additivity, Utility, and Subjective Probability," *Journal of Mathematical Psychology* 4, 175-201.

Yaari, M. E. (1987). "The Dual Theory of Choice Under Risk," *Econometrica* 55, 95-115.

Wakker, P. P. (1989). *Additive Representations of Preferences.* Boston: Kluwer.