1.2

Decisions Without Ordering

TEDDY SEIDENFELD, MARK J. SCHERVISH,
AND JOSEPH B. KADANE

ABSTRACT

We review the axiomatic foundations of subjective utility theory with a view toward understanding the implications of each axiom. We consider three different approaches, namely, the construction of utilities in the presence of canonical probabilities, the construction of probabilities in the presence of utilities, and the simultaneous construction of both probabilities and utilities. We focus attention on the axioms of independence and weak ordering. The independence axiom is seen to be necessary to prevent a form of Dutch Book in sequential problems.

Our main focus is to examine the implications of not requiring the weak order axiom. We assume that gambles are partially ordered. We consider both the construction of probabilities when utilities are given and the construction of utilities in the presence of canonical probabilities. In the first case we find that a partially ordered set of gambles leads to a set of probabilities with respect to which the expected utility of a preferred gamble is higher than that of a dispreferred gamble. We illustrate some comparisons with theories of upper and lower probabilities. In the second case, we find that a partially ordered set of gambles leads to a set of lexicographic utilities, each of which ranks preferred gambles higher than dispreferred gambles.

I. INTRODUCTION: SUBJECTIVE EXPECTED UTILITY [SEU] THEORY

The theory of (subjective) expected utility is a normative account of rational decision making under uncertainty. Its well-known tenets are

Reprinted from W. Sieg (ed.), *Acting and Reflecting: The Interdisciplinary Turn in Philosophy* (Dordrecht: Kluwer Academic Publishers, 1990), 143–170.

	s_1	■	■	■	s_j	■	■	■	s_n
A_1	o_{11}				o_{1j}				o_{1n}
■									
A_i	o_{i1}				o_{ij}				o_{in}
■									
A_m	o_{m1}				o_{mj}				o_{mn}

Figure 1. Canonical decision matrix

spotlighted by the familiar, canonical decision problem in which s_j: $j = 1, \ldots, n$ is a partition, and o_{ij} is the outcome of option$_i$ (act$_i$) in state$_j$. That is, acts are functions from states to outcomes. This problem is illustrated in Figure 1.

In the canonical decision problem, states are *value-neutral* and *act independent*. The value of an outcome does not depend upon the state in which it is rewarded, and the choice of an act does not alter the agent's opinion (uncertainty) about the states. In insurance terms, there are no "moral hazards."

General Assumption. Acts are weakly ordered by (weak) preference, $\preceq$, a reflexive, transitive relation with full comparability between any two acts.

Subjective Expected Utility [SEU] Thesis. There is a real-valued utility $U(\ldots)$, defined over outcomes, and a personal probability $p(\ldots)$, defined over states, such that

$$A_1 \preceq A_2 \quad \text{if and only if} \quad \sum_j p(s_j)U(o_{1j}) \leq \sum_j p(s_j)U(o_{2j}).$$

There are several well-trodden approaches to the normative justification of the SEU thesis, which we discuss in the remainder of this section.

I.1. *Utility Given Probability*

The seminal efforts of J. von Neumann and O. Morgenstern (1947) provide necessary and sufficient conditions for an expected utility representation of preference over (simple) *lotteries*: acts specified by a probability on a (finite subset of a) set of rewards. Their theory uses one "structural" axiom and three axioms on preference $\preceq$.

Structural Axiom. Acts are simple lotteries (L_i), i.e., simple distributions over a set of rewards. The domain of acts is closed under convex combinations of distributions – denoted by $\alpha L_1 + (1-\alpha)L_2$.

Weak-Order Axiom. $\preceq$ is a reflexive, transitive relation over pairs of lotteries, with comparability between any two lotteries.

Independence Axiom. For all L_1, L_2, L_3 $(0 < \alpha \le 1)$,

$$L_1 \preceq L_2 \text{ if and only if } \alpha L_1 + (1-\alpha)L_3 \preceq \alpha L_2 + (1-\alpha)L_3.$$

Archimedean Axiom. For all $(L_1 \prec L_2 \prec L_3)$ $\exists(0 < \alpha, \beta < 1)$,

$$\beta L_1 + (1-\beta)L_3 \prec L_2 \prec \alpha L_1 + (1-\alpha)L_3.$$

A particularly simple illustration of this theory involves lotteries over three rewards $(r_1 \prec r_2 \prec r_3)$, where the reward r_i is identified with the degenerate lottery having point-mass $P(r_i) = 1$ $(i = 1, 2, 3)$. Following the excellent presentation by Machina (1982), we have a simple geometric model for what is permitted by expected utility theory. Figure 2 depicts the consequences of the axioms.

According to the axioms, indifference curves ($\sim$) over lotteries are parallel, straight lines of (finite) positive slope. L_i is (strictly) preferred to L_j, $L_j \prec L_i$, just in case the indifference curve for L_i is to the left of the indifference curve for L_j. Hence, in this setting, expected utility theory permits one degree of freedom for preferences, corresponding to the choice of a slope for the lines of indifference.

Another version of this example occurs with the decision theoretic reconstruction of "most powerful" Neyman-Pearson tests of a simple "null" hypothesis (h_0) versus a simple rival alternative (h_1). We face the binary decision given by the matrix:

	h_0	h_1
accept h_0	a	b
reject h_0	c	d

where we suppose that outcomes b and c are each dispreferred to either outcomes a and d. In the usual jargon, c is the outcome of a type$_1$

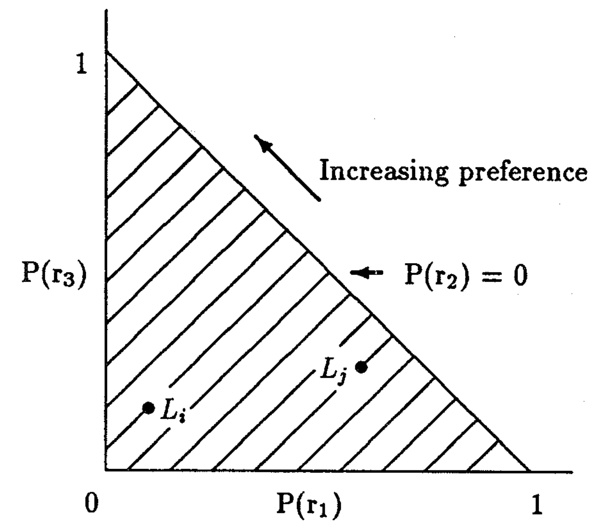


Figure 2. Curves of indifference with three rewards

error and b is the outcome of a type$_2$ error. By the assumption that states are "act independent," without loss of generality, we may rewrite the matrix with utility outcomes:

	h_0	h_1
accept h_0	0	$-(1-x)$
reject h_0	$-x$	0

where $0 < x < 1$. The expected utility hypothesis requires that accepting h_0 is not preferred to ($\preceq$) rejecting h_0 just in case $(1 - p_0)/p_0 \ge x/(1-x)$, where p_0 is the "prior" probability of h_0.

Suppose we have the option of conducting an experiment E (with a sample space of possible experimental outcomes denoted by Ω), where the conditional probabilities $p(\cdot|h_0)$ and $p(\cdot|h_1)$ over Ω are specified by the description of E. A (Neyman-Pearson) statistical test of h_0 against h_1, based on E, is defined by a critical region $\mathcal{R} \subset \Omega$; with the understanding that h_0 is rejected iff $\mathcal{R}$ occurs. Associated with each statistical test are two quantities: (α, β), where $\alpha = p(\mathcal{R}|h_0)$ is the probability of a type$_1$ error, and $\beta = p(\mathcal{R}^c|h_1)$ is the probability of a type$_2$ error.

According to the N-P theory, two tests may be compared by their (α, β) numbers. Say that T_2 dominates T_1 if $(\alpha_2 \le \alpha_1)$, $(\beta_2 \le \beta_1)$ and at least one of these inequalities is strict. This agrees with the ranking of

Table 1. *The "best" β-values for twelve α-values and six experiments*

σ =	.250	.333	.400	.500	1.000	1.333
α	β-values					
.010	.047	.250	.431	.628	.908	.942
.020	.026	.172	.327	.521	.854	.904
.030	.017	.131	.268	.452	.811	.871
.040	.012	.106	.227	.401	.773	.841
.045	.011	.096	.210	.380	.756	.828
.050	.009	.088	.196	.361	.740	.814
.055	.008	.080	.184	.344	.725	.802
.060	.007	.074	.172	.328	.710	.789
.070	.006	.064	.153	.300	.683	.766
.080	.005	.055	.137	.276	.657	.744
.090	.004	.049	.123	.255	.633	.722
.100	.003	.043	.111	.236	.611	.702

tests by their expected utility since (prior to observing the outcome of the experiment) the expected utility of test T, having errors (α, β), is given by:

$$-[x \cdot p(\mathcal{R}\&h_0) + (1-x) \cdot p(\mathcal{R}^c\&h_1)] = -[x \cdot \alpha \cdot p_0 + (1-x) \cdot \beta \cdot (1-p_0)],$$

so that $T_1 < T_2$ if T_2 dominates T_1 (except for the trivial cases of certainty: $p_0 = 0$ or $p_0 = 1$, when $T_2 \sim T_1$ is possible still – but then there hardly is need for a "test" of h_0).

Given an experiment E, there are numerous, mutually undominated tests based on E. For example, consider the family of undominated tests of h_0: $\mu = 0$ versus h_1: $\mu = 1$ from the observation of a normally distributed random variable $X \sim N[\mu, \sigma^2]$, with specified variance σ^2. These are just the family of "best," i.e., most powerful tests of h_0 versus h_1 – which, by the Neyman-Pearson lemma, is the family of likelihood ratio tests for the datum x. Table 1 lists some (α, β) values for undominated tests from six such experiments: $\sigma = 1/4$; $= 1/3$; $= 2/5$; $= 1/2$; $= 1$; and $= 4/3$.

Three of these families, corresponding to $\sigma = 1/3$; $\sigma = 1/2$; and $\sigma = 4/3$, are depicted by the curves in Figure 3. The graph shows the tangents to these three curves at $\alpha = 0.05$. The "0.05-α-level" tangents are

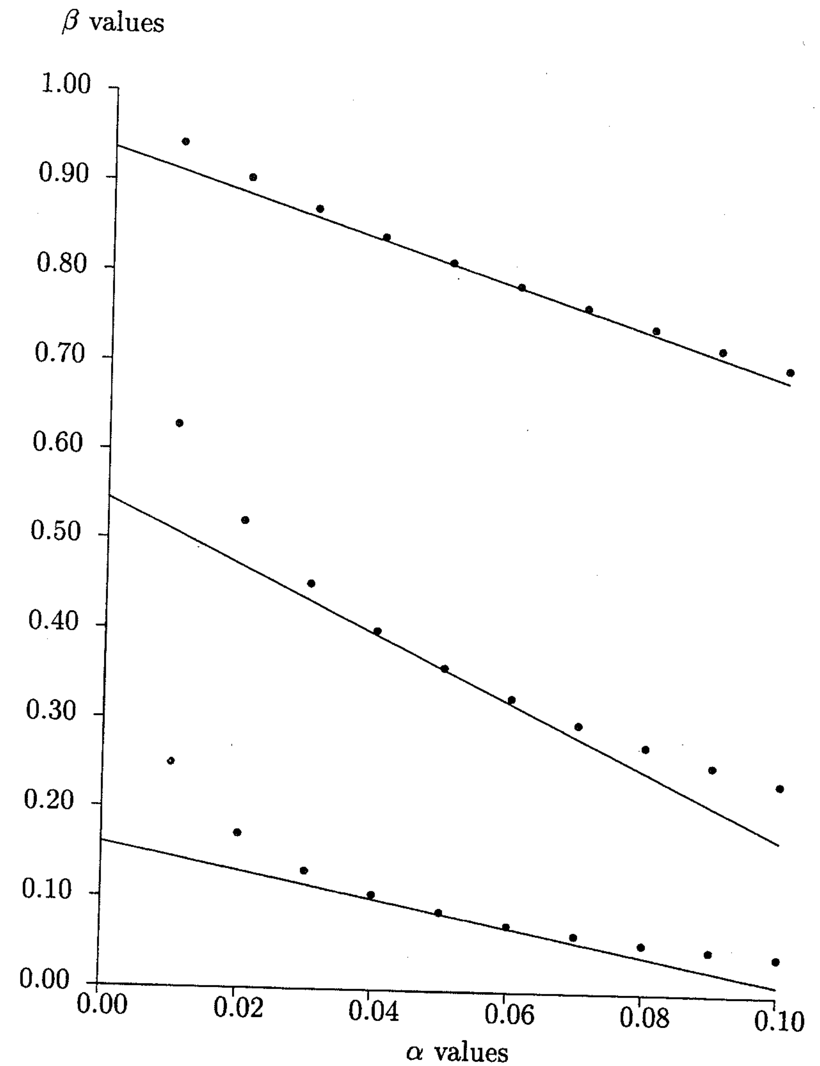


Figure 3. Families of (α, β) pairs for undominated tests

not parallel. A statistical test of h_0 versus h_1 is a lottery involving the three prizes $-x$, $-(1-x)$, 0. As before, if the preferences among such tests satisfy the expected utility hypothesis, then the indifference curves (of equally desirable tests) are parallel straight lines.

In Figure 3, these indifference curves have negative slopes equal to $-xp_0/(1-x)(1-p_0)$. (The slopes are negative because *smaller* (α, β)

values are better.) Thus, expected utility theory is in conflict with the popular *convention* of choosing the "best" test with a fixed α-level, e.g., $\alpha = 0.01$ or $\alpha = 0.05$. That is, when testing simple hypotheses, in order to agree with expected utility theory the choice of α must reflect the precision of the experiment. (See also Lindley [1972, p. 14], who gives this argument for the special case of "0-1" losses.) In a purely "inferential" (non-decision-theoretic) Bayesian treatment of testing a simple hypothesis versus a composite alternative, Jeffreys (1971, p. 248) argues for the same caveat about constant α-levels.

A dramatic illustration of this lesson can be seen with the aid of Table 1. Suppose an agent prefers undominated tests with $\alpha = 0.05$ over rivals. Then, for the experiment corresponding to $\sigma = 1/4$, test T_2 is preferred to test T_1, where ($\alpha_1 = 0.01, \beta_1 = 0.047$) and ($\alpha_2 = 0.05, \beta_2 = 0.009$). Likewise, for the experiment corresponding to $\sigma = 4/3$, test T_4 is preferred to test T_3, where ($\alpha_3 = 0.09, \beta_3 = 0.722$) and ($\alpha_4 = 0.05, \beta_4 = 0.814$). However, test T_5, the "50-50" mixture of tests T_1 and T_3, is preferred to test T_6, the "50-50" mixture of tests T_2 and T_4, as ($\alpha_5 = 0.05, \beta_5 = 0.385$) and ($\alpha_6 = 0.05, \beta_6 = 0.412$), so that T_5 dominates T_6. This is the decision-theoretic analogue of Cox's (1958) example involving the failure of the ancillarity principle within Neyman-Pearson theory.

I.2. *Probability Given Utility*

The "Dutch Book" argument, tracing back to Ramsey (1931) and deFinetti (1937), offers prudential grounds for action in conformity with personal probability. Under several "structural" assumptions about combinations of stakes (that is, assumptions about the combination of wagers), your betting policy is consistent ("coherent") only if your "fair" odds are probabilities.

A simple bet on/against event E, at odds of $r: 1 - r$, with a total stake $S > 0$ (say, bets are in $ units), is specified by its payoffs, as follows:

	E	¬E
bet on E	win $(1 - r)S$	lose rS
bet against E	lose $(1 - r)S$	win rS

(By writing $S < 0$ we can reverse betting "on" or "against.")

The general assumption (that acts are weakly ordered by $\preceq$) entails

that there is a preference among the options betting on, betting against and abstaining from betting (whose consequences are "status quo," or net $0, regardless of whether E or ¬E). The special ("structural") assumptions about the stakes for bets require, in addition:

a. Given an event E, a betting rate $r: 1 - r$ and a stake S, your preferences satisfy exactly one of three profiles. Either:
   betting on $\prec$ abstaining $\prec$ betting against E,
   or betting on $\sim$ abstaining $\sim$ betting against E,
   or betting against $\prec$ abstaining $\prec$ betting on E.
b. The (finite) conjunction of favorable/fair/unfavorable bets is favorable/fair/unfavorable. (A conjunction of bets is favorable in case it is preferred to abstaining, unfavorable if dispreferred to abstaining, and fair if indifferent to abstaining.)
c. Your preference for outcomes is continuous in rates; in particular, each event E carries a unique "fair odds" r_E for betting on E.

Note: It follows from these assumptions that your attitude towards a simple bet is independent of the size of the stake.

Dutch Book Theorem. If your fair betting odds are not probabilities, then your preferences are incoherent, i.e., inconsistent with the preference for sure-gains. Specifically, then there is some "favorable" combination of bets which is dominated by abstaining, i.e., some "favorable" combination where you pay out in each state of a finite (exhaustive) partition. (See Shimony (1955), for an elegant proof using the linear structure of these bets.)

The Dutch Book argument can be extended to include conditional probability, $p(\cdot|\cdot)$, through the device of called-off bets. A called-off bet on (against) H given E, at odds of $r: (1 - r)$ with total stake S (>0), is specified by its payoffs, as follows.

	H ∩ E	¬H ∩ E	¬E
bet on H	win $(1 - r)S$	lose rS	0 (the bet is called off)
bet against H	lose $(1 - r)S$	win rS	0 (the bet is called off)

By including called-off bets within the domain of act to be judged favorable/indifferent/unfavorable against abstaining, and subject to the same structural assumptions (a–c) imposed above, coherence of "fair"

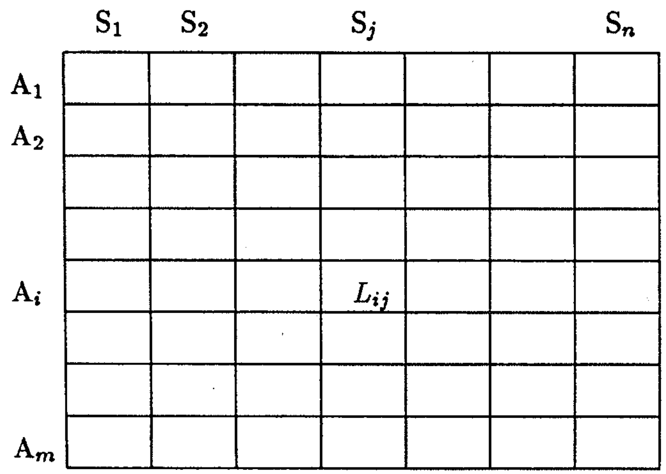


Figure 4. Anscombe-Aumann "horse lotteries"

betting odds entails: $r_{(H|E)} \cdot r_E = r_{(H \cap E)}$, where "$r_{(H|E)}$" is the "fair called-off" odds on H given E. This result gives the basis for interpreting conditional probability, $p(H|E)$, by the fair "called-off:" odds $r_{(H|E)}$, for then we have:

$$p(H|E) \cdot p(E) = p(H \cap E),$$

the axiomatic requirement for conditional probabilities.

I.3. *Simultaneous Axiomatizations of (Personal) Probability and Utility*

We distinguish two varieties:

i. without extraneous "chances," as in Savage's (1954) theory.
ii. with extraneous "chances," a continuation of the von Neumann-Morgenstern approach, as in Anscombe & Aumann's (1963) theory of "horse lotteries." Horse lotteries are a generalization of lotteries, as illustrated in Figure 4.

An outcome of act A_i, when state S_j obtains (when "horse$_j$" wins), is the von Neumann-Morgenstern lottery L_{ij}. The Anscombe-Aumann theory is the result of taking the von Neumann-Morgenstern axiomatization of $\preceq$ (the Weak-order, Independence and Archimedean postulates), and adding an assumption that states are value-neutral.

II. INDEPENDENCE AND CONSISTENCY IN SEQUENTIAL CHOICES

We are interested in relaxing the "ordering" postulate, without abandoning the normative standard of coherence (consistency) and without losing the representation ("measurement") of our modified theory. First, however, let us compare two programs for generalizing expected utility in order to justify the concern for consistency:

Program $\neg$***I*** *– Delete the "Independence" Postulate.* Illustrations: Samuelson (1950); Kahneman & Tversky's "Prospect Theory" (1979); Allais (1979); Fishburn (1981); Chew & Macrimmon (1979); McClennen (1983); and especially Machina (1982, 1983 – which has an extensive bibliography).

Program $\neg$***O*** *– Delete the "Ordering" Postulate.* Illustrations: I. J. Good (1952); C. A. B. Smith (1961) – related to the Dutch Book argument; I. Levi (1974, 1980); Suppes (1974); Walley & Fine (1979); Wolfenson & Fine (1982); Schick (1984).

And in Group Decisions: Savage (1954, §7.2); Kadane & Sedransk (1980); and Kadane (1996) – applied to clinical trials.

Also, "regret" models involve a failure of "ordering" if we define the relation $\preceq$ by their choice functions, which violate (Sen's properties α and β, 1977) "independence of irrelevant alternatives": Savage (1954, §13.5); Bell & Raiffa (1979); Loomes & Sugden (1982); and Fishburn (1983).

A CRITICISM OF PROGRAM ¬I. Consider elementary problems where we apply the modified theory ¬**I** to simple lotteries. Thus, we discuss the case, like the von Neumann-Morgenstern setting, where "probability" is given and we try to quantify (represent) the value of "rewards."

There is a technical difficulty with the theory that results from just the two postulates of "weak-ordering" and the usual "Archimedean" requirement. It is that these two are insufficient to guarantee a real-valued "utility" representation of $\preceq$ (see Fishburn, 1970, §3.1). We can avoid this detail and also simplify our discussion by assuming that lotteries are over (continuous) monetary rewards; we assume that lotteries have $-equivalents and more $ is better.

Under these assumptions and to underscore the normative status of coherence, let us investigate what happens when a particular consequence of "independence" is denied.

Mixture Dominance ("Betweenness"). If lotteries L_1, L_2 are each preferred (dispreferred) to a lottery L_3, so too each convex combination of L_1 and L_2 is preferred (dispreferred) to L_3.

Here is an illustration of sequential inconsistency for a failure of mixture dominance. Let $L_1 \sim L_2 \sim \$5.00$, but $0.5L_1 + 0.5L_2 \sim \$6.00$: the agent prefers the "50-50" mixture of L_1 and L_2 to each of them separately. Then, by continuity of (ordinal) utility over dollar payoffs, there is a fee, $-\$\ \varepsilon$, such that, e.g.,

$$L_1 \sim L_2 \prec 0.5(L_1 - \varepsilon) + 0.5(L_2 - \varepsilon) \sim \$5.75 \prec 0.5L_1 + 0.5L_2,$$

where $L_i - \varepsilon$ denotes the modification of L_i obtained by reducing each payoff in L_i by the fee $\$\ \varepsilon$. Assume $\$4.00 \prec (L_i - \varepsilon)(i = 1, 2)$.

Consider two versions of a sequential decision problem, depicted by the decision trees in Figures 5 and 6. "Choice" nodes are denoted by a □ and "chance" nodes are denoted by •. In the first version (Figure 5), at node **A** the agent may choose between plans **1** and **2**. These lead to *terminal* choices at nodes **B**, depending upon how a "fair" coin lands at the intervening chance nodes. If the agent chooses plan **1** (at **A**) and the coin lands "heads," he faces a (terminal) choice between lottery L_1 and the certain prize of \$5.50. If, instead, the coin lands "tails," he faces a (terminal) choice between L_2 and the certain prize of \$5.50.

The decision tree is known to the agent in advance. He can anticipate (at **A**) how he will choose at subsequent nodes, if only he knows what his preferences will be at those junctures. In the problem at hand, we suppose the agent knows that, at **B**, he will not change his preferences over lotteries. (There is nothing in the flip of the coin to warrant a shift in his valuation of specified, von Neumann-Morgenstern lotteries.) For example, according to our assumptions, at **A** he prefers a certain \$5.50 to the lottery L_1. Thus, we assume that at **B**, too, he prefers the \$5.50 to L_1.

Then, at **A**, the agent knows which terminal options he will choose at nodes **B** and plans accordingly. If he selects plan **1**, he will get \$5.50. If he selects plan **2**, he will get lottery $L_1 - \varepsilon$ with probability 1/2 and he will get lottery $L_2 - \varepsilon$ with probability 1/2. But this he values \$5.75; hence, plan **2** is adopted.

The decision program **¬I** requires the "ordering" postulate for terminal decisions. Thus, at choice nodes such as **B**, the agent is indifferent between lotteries that are judged equally desirable (~) according to his preferences (⪯). The second version of the sequential choice problem (Figure 6) results by replacing the lotteries at the (terminal)

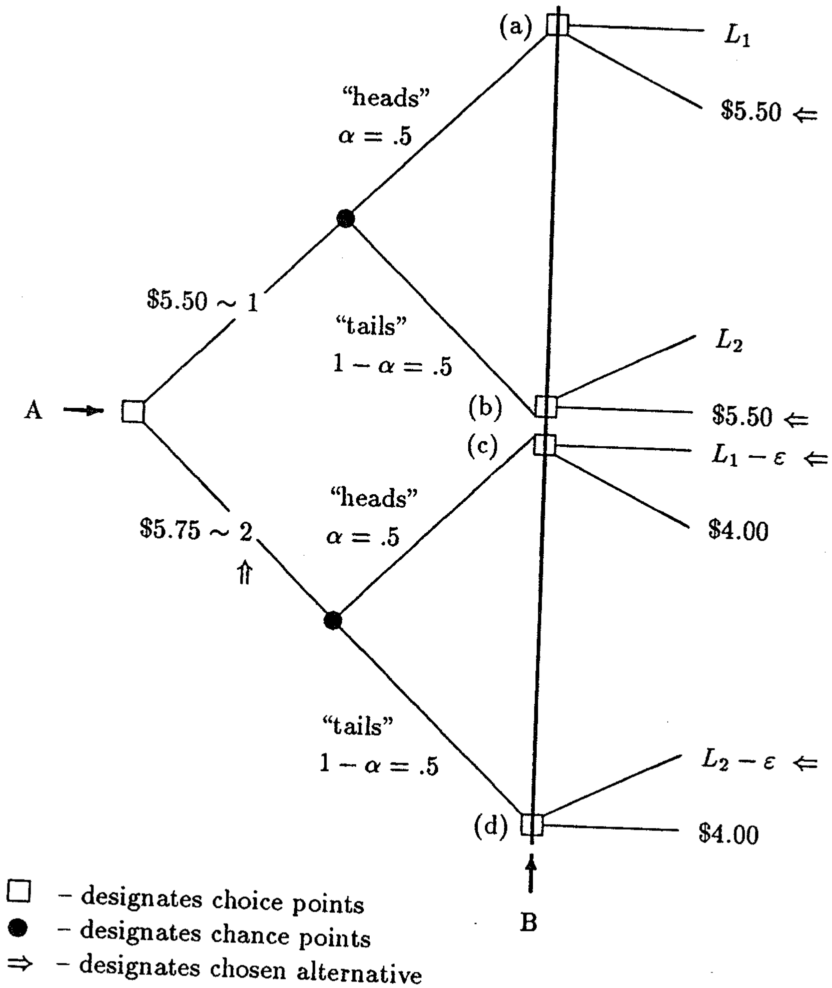


Figure 5. First version of the sequential decision: an illustration of sequential incoherence for a failure of mixture dominance ("betweenness"). At choice node A option 2 is preferred to option 1. At each choice node B this preference is reversed.

nodes **B** by their sure-dollar equivalents under ~. In this version, by the same reasoning, the agent rejects plan **2** and adopts plan **1**. This is an inconsistency within the program since, at nodes **B**, the agent's preferences are given by the weak-ordering, ⪯, yet his (sequential) choices do not respect the indifferences, ~, generated by ⪯.

Let us call such inconsistency in sequential decisions an episode of

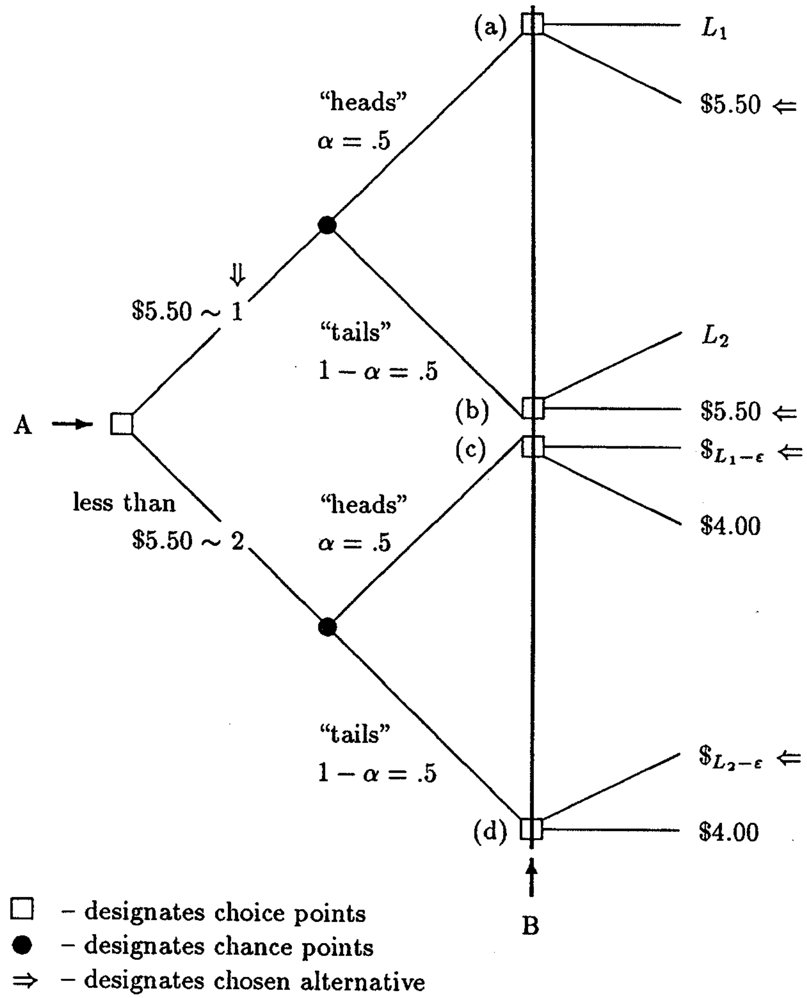


Figure 6. Second version of the sequential decision: an illustration of sequential incoherence for a failure of mixture dominance ("betweenness"). At choice node A option 1 is preferred to option 2. The tree results by replacing $L_i - \varepsilon$ $(i = 1, 2)$ from Figure 6.5 with \$-equivalents under $\preceq$.

"sequential incoherence." Then, we can generalize this example and show:

Theorem. If $\preceq$ is a weak order **(1)** of simple lotteries satisfying the Archimedean postulate **(3)** with sure-dollar equivalents for lotteries, and if $\preceq$ respects stochastic dominance in payoffs (a greater chance at

more \$ is better), then a failure of "independence" **(2)** entails an episode of sequential incoherence (see Seidenfeld, 1988).

However, as Levi's decision theory – one which relaxes the ordering postulate rather than "independence" – avoids sequential incoherence (Levi, 1986), we see that it is not necessary for decisions to agree with expected utility theory in order that they be sequentially coherent.

III. REPRESENTATION OF PREFERENCES WITHOUT "ORDERING"

Next, we discuss the representation of an alternative theory falling within program $\neg$**O**: to weaken the "ordering" assumption. Again, let us begin with the more elementary problem where we try to quantify values for the rewards when "probability" is given – analogous to the von Neumann-Morgenstern setting.

Let $R = \{r_i: i = 1, \ldots\}$ be a countable set of rewards, and let $L = \{L: L$ is a discrete lottery, a discrete P on $R\}$. As before, define the convex combination of two lotteries $\alpha L_1 + (1 - \alpha)L_2 = L_3$, by $P_3 = \alpha P_1 + (1 - \alpha)P_2$. We consider a theory with three axioms:

Axiom 1. Preference $\prec$ is a strict partial order, being transitive and irreflexive. (This weakens the "weak order" assumption, since non-comparability, ~, need not be transitive.)

Axiom 2. (independence). For all L_1, L_2, and L_3, and for all $1 \geq \alpha > 0$,

$$L_1 \prec L_2 \text{ iff } \alpha L_1 + (1-\alpha)L_3 \prec \alpha L_2 + (1-\alpha)L_3.$$

Axiom 3. A suitable Archimedean requirement. (Difficulties with axiom 3 are discussed below.)

Say that a real-valued utility U *agrees with* the partial order $\prec$ iff

$$\sum_i P_1(r_i)U(r_i) < \sum_i P_2(r_i)U(r_i) \text{ whenever } L_1 \prec L_2.$$

We hope to show that $\prec$ is represented by a (convex) set of agreeing utilities. That is, we seek to show there is a (maximal) set of agreeing utilities, $\mathbf{U}\prec$, where (by the **unanimity** rule)

$L_1 \prec L_2$ **iff** for each $U \in \mathbf{U} \prec \sum_i P_1(r_i)U(r_i) < \sum_i P_2(r_i)U(r_i)$.

Aside on Related Results. Aumann (1962) proved that when R is finite, there exists a real-valued utility agreeing with $\prec$, provided axioms like 1–3 hold. A lottery is *simple* if its support is a finite set of rewards. Kannai (1963) extended Aumann's result to simple lotteries on a countable reward set by strengthening the Archimedean axiom 3. (More precisely, these theories deal with a *reflexive* and transitive partial order – which identifies indifferences – not just with the irreflexive part $\prec$.) These two studies, as well as Fishburn's (1970, ch. 9) simplification of Aumann's work, use an embedding of the partial order in a separable, normed linear space. Their proofs have a common theme. Represent a lottery L by a vector of its probability P, with coordinates corresponding to the elements of R. Because a lottery is simple, all but finitely many of its coordinates are zero. Call a vector difference $(P_2 - P_1)$ "favorable" when $L_1 \prec L_2$. The set of "favorable" vectors forms a convex cone, and a Separating Hyperplane Theorem (Klee, 1955) yields a utility. (However, the separability assumption prohibits using this method when, e.g., the reward set R is uncountable.)

There are three observations which help to explain some of the difficulties that arise in carrying out our project for representing preferences given by partial orders.

1. The usual Archimedean axiom won't do; it is too restrictive.

Example 1. $R = \{r_0 \prec r^* \prec r_1\}$ but for no $0 < \alpha < 1$ is it the case that $\alpha r_0 + (1 - \alpha)r_1 \prec r^*$. However, this partial order can be represented by a set of utilities, $\mathbf{U} = \{U_x: 0 < x < 1\}$ with $U_x(r_0) = 0$, $U_x(r_1) = 1$ and $U_x(r^*) = x$. This is illustrated in Figure 7.

Hence, in general, to represent a partial order generated by a set of utilities, a weakening of the usual Archimedean postulate is necessary.

2. Two different convex sets of utilities can generate the same partial order. That is, given convex sets $\mathbf{U}_1$ and $\mathbf{U}_2$, we can difine the partial orders $\prec_1$ and $\prec_2$ according to the "unanimity" rule. However,

Example 2. It may be that $\prec_1 = \prec_2$, though $\mathbf{U}_1 \neq \mathbf{U}_2$. See Figure 8 for an illustration.

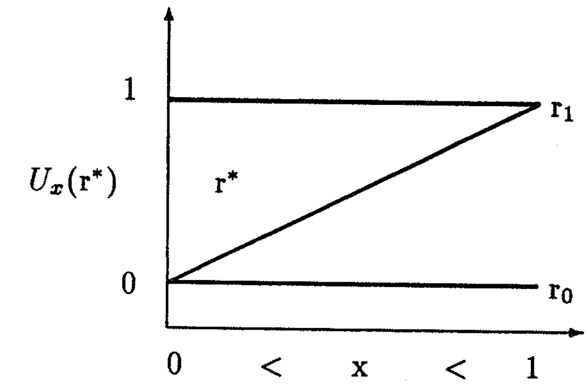
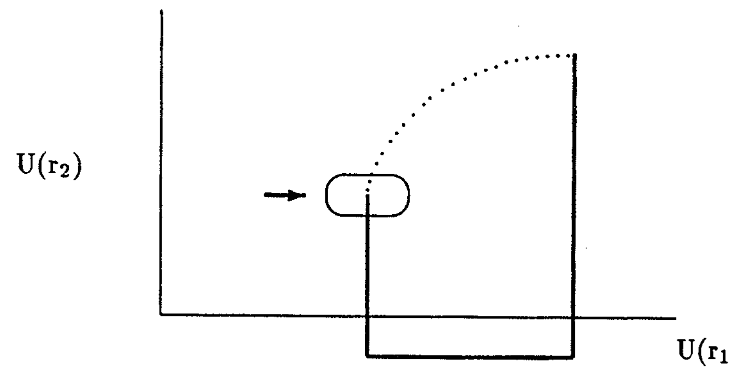


Figure 7. Example of restrictions of the usual Archimedean axiom



··· designates an open boundary
— designates a closed boundary

Figure 8. Two convex sets of utilities which generate the same partial order. The two (convex) sets differ by the presence of the point identified by the arrow. The common partial order is generated by the "unanimity" rule

When we shift from representing indeterminate utility (given determinate "chances") to the dual task of representing indeterminate probability (given a determinate utility – by assuming favorable bets combine according to the Dutch Book assumptions – see §IV), this phenomenon causes difficulties for the representation of conditional probabilities. (Also, contrast this with Aumann's example, 1964, p. 210.)

3. Last, though the representation of indeterminate preferences over lotteries (given determinate "chances") is by convex sets of utilities – similarly the dualized representation of indeterminate betting odds (given bets are in stakes which behave like utiles – see §IV) is by convex sets of probabilities – when we turn to the simultaneous representation of indeterminate preferences and beliefs (through "horse lotteries"), convexity may fail. The set {(P,U)} of probability-utility pairs which agree with a partially ordered preference over horse lotteries may not be convex (nor even connected). However, convexity is assured for both sets: {(P, U*): U* fixed} and {(P*, U): P* fixed}.

Here is an example of non-convexity of the set of probability-utility pairs agreeing with a partial order, $\prec$, over "horse lotteries."

Example 3. There are two uncertain states $(S, \neg S)$ and three rewards (r_0, r^*, r_1), with r_1 preferred to r_0, $r_0 \prec r_1$, but where r^* is $\prec$ – incomparable with either r_0 or r_1. Consider the two acts, $A1$ and $A2$, defined by the payoffs:

	S	$\neg S$
$A1$	r_0	r_1
$A2$	r_1	r^*

Fix the utilities $U(r_0) = 0$ and $U(r_1) = 1$, and let $P(S)$ denote the probability of state S. Then Figure 9 shows the regions where $A1$ is preferred or $A2$ is preferred.

This example shows why the proof techniques based on the Separating Hyperplane results are inappropriate for identifying the (maximal) set of pairs: {(P, U): (P, U) agrees with $\prec$} for "horse lotteries."

Our proof procedure for giving a representation of a strict preference over horse lotteries is to modify Szpilrajn's (1930) argument that, by transfinite induction, every partial order may be extended to a total order. The modification involves preserving the other axioms: "Independence," "Archimedes," and "value-neutrality" of states. In the Appendix we illustrate this technique for representing strict partial orders of von Neumann-Morgenstern lotteries by convex sets of (lexicographic) utilities.

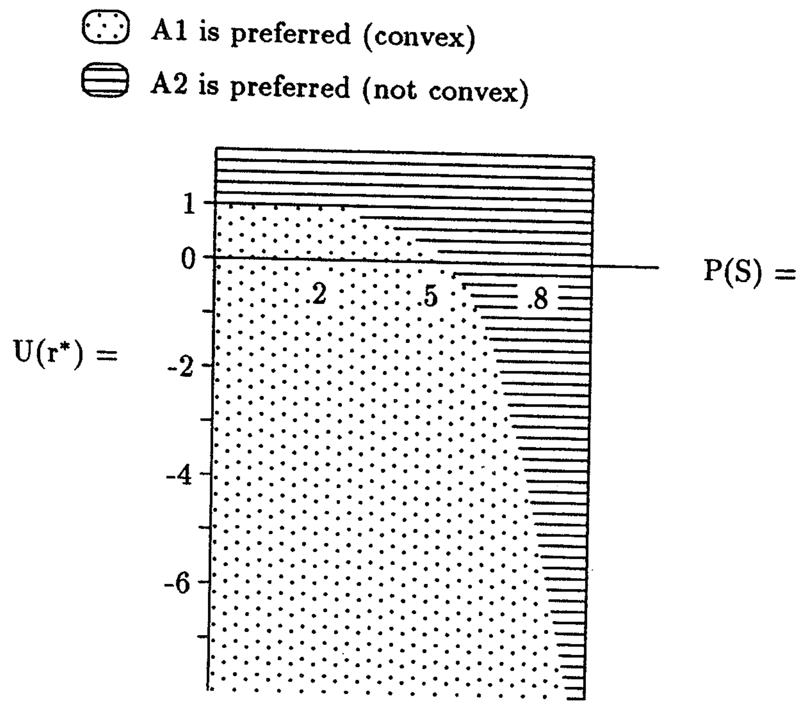


Figure 9. Regions of preference for Example 3

IV. REPRESENTATION OF BELIEFS WITHOUT "ORDERING"

By appeal to the Separating Hyperplanes theorem, we may generalize the Dutch Book argument to establish the coherence of beliefs for partially ordered gambles, including the case (discussed by C. A. B. Smith, 1961) of "medial" odds. Consider the finite partition of states $\{s_j: j = 1, \ldots, n\}$, and define a gamble as a vector of n real-values, $A_i = \langle r_{i1}, \ldots, r_{in} \rangle$, where r_{ij} is the (utility of the) reward generated by A_i when state s_j obtains. Denote the constant gamble $r_j = 0$ (corresponding to "no bet," or "status quo") by **O**, and define the set of *favorable gambles*, $\mathcal{F}$, to be those which are preferred to **O** in pairwise comparisons. As in the Dutch Book argument, we make structural assumptions about the value of the rewards, assuring that the magnitudes of the rewards behave like utilities.

STRUCTURAL ASSUMPTIONS

i. Weak dominance over **O**. If $r_{ij} \geq 0$ (all j) with a strict inequality for some j, then A_i is favorable.

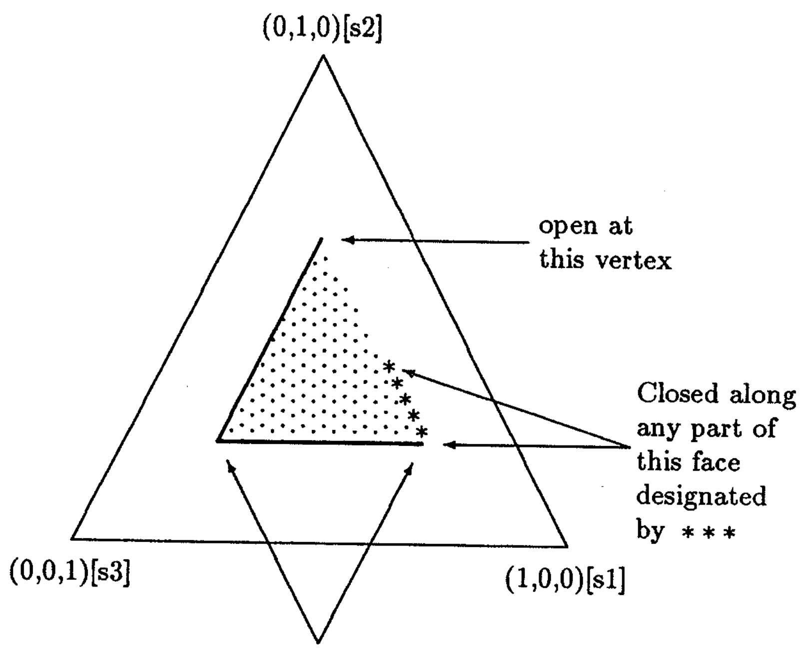


Figure 10. Different convex sets of probabilities which generate the same partial order under the "unanimity" rule

ii. Scalars. If A_i is favorable, so too is $cA_i = \langle \ldots, cr_{ij}, \ldots \rangle$, for $c > 0$.
iii. Convex combinations. If A_h and A_i are favorable, so too is the convex combination $xA_h + (1-x)A_i = \langle \ldots, xr_{hj} + (1-x)r_{ij}, \ldots \rangle$, for $0 \le x \le 1$.

REPRESENTATION THEOREMS RELATING TO $\mathcal{F}$

Theorem 1. Coherence of $\mathcal{F}$:

i. $\mathbf{O} \notin \mathcal{F}$ iff there is a maximal, non-empty convex set $\mathcal{P}$ of probabilities with the property that $\forall A_i \in \mathcal{F}, \forall p \in \mathcal{P}, \Sigma_j p(s_j) r_{ij} > 0$.
ii. Moreover, if $\mathcal{F}$ is open, or if $\mathcal{F} \cup \{\mathbf{O}\}$ is closed, then $A_i \in \mathcal{F}$ provided $\forall p \in \mathcal{P}, \Sigma_j p(s_j) r_{ij} > 0$.

We may extend this to include conditional probabilities by paralleling the device of "called-off" bets, used to show coherence of conditional odds in the Dutch Book argument. Then:

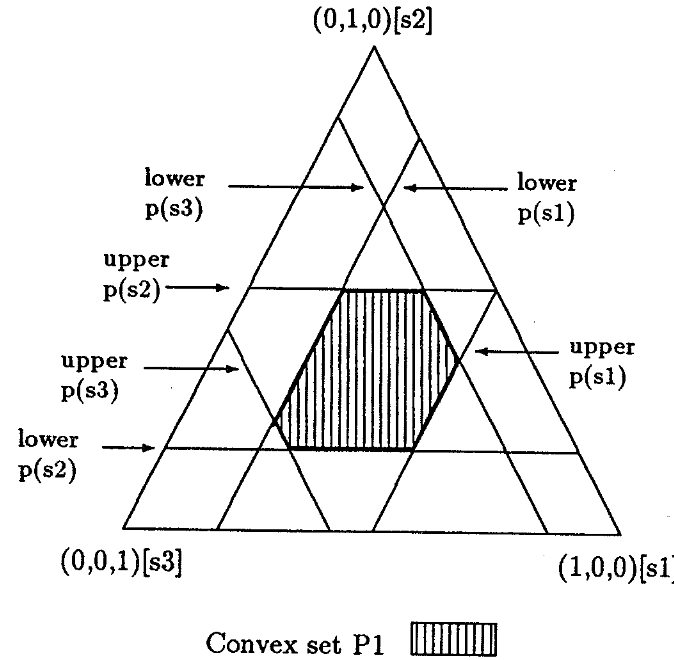


Figure 11. Supporting lines determined by odds alone

Theorem 2. Coherence of conditionally favorable gambles: Let $\mathcal{F}_E (\subset \mathcal{F})$ be the set of favorable gambles, called off in case event E fails to occur, i.e., $\forall A_i \in \mathcal{F}_E, r_{ij} = 0$ if $s_j \in E^C$. Assume that $\mathbf{O} \notin \mathcal{F}$.

i. Then $\forall A_i \in \mathcal{F}_E, \forall p \in \mathcal{P}, \Sigma_j p(s_j|E) r_{ij} > 0$.
ii. Moreover, if A_i is called off when E fails and either $\mathcal{F}_E$ is open or $\mathcal{F}_E \cup \{\mathbf{O}\}$ is closed, then $A_i \in \mathcal{F}_E$ provided $\forall p \in \mathcal{P}, \Sigma_j p(s_j|E) r_{ij} > 0$.

In both theorems, the closure conditions imposed in clause (ii) reflect the severity of the problem illustrated in Figure 10, which is dual to the problem illustrated in Example 2, p. 54.

The favorable gambles $\mathcal{F}$ are a subset of those preferred to "no bet" under the partial order ($\prec_{\mathcal{P}}$), generated by the "unanimity" rule adapted to the set $\mathcal{P}$. Denote the closure of $\mathcal{F}$ by cl($\mathcal{F}$), and denote by $\mathcal{F}^-$ the set that results from taking each favorable gamble and changing the sign of its payoffs. It is straightforward to verify that $\mathcal{P}$ is a unit set (expected utility theory) just in case $\prec_{\mathcal{P}}$ is a weak-order. That occurs if and only if cl($\mathcal{F}$) $\cup \mathcal{F}^- = \mathcal{R}^n$ (the space of all gambles on the n states s_j). In other words, when $\mathcal{P}$ is not a unit set, there will be

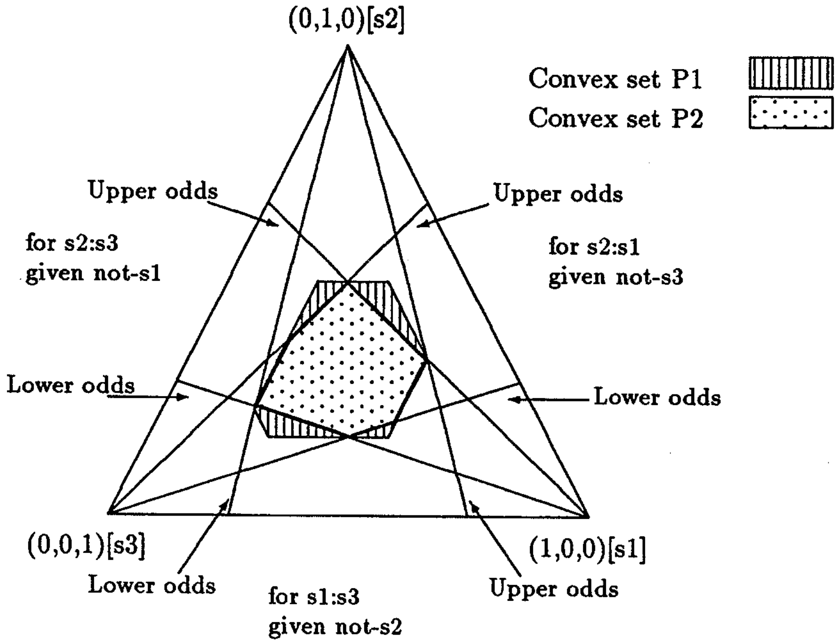


Figure 12. Supporting lines determined by odds and called-off bets

gambles A_1 and A_2 with $A_1 \prec_{\mathcal{P}} A_2$ but where none of A_1, $A_{\bar{1}}$, A_2, or $A_{\bar{2}}$ is favorable.

We illustrate sets $\mathcal{P}$ for the elementary case of three states, $n = 3$ in Figures 11–13. The figures use barycentric coordinates. Each trinomial distribution on $\{s_1,s_2,s_3\}$ is a point in the simplex having vertices: ⟨(100) (010) (001)⟩. Figure 10 shows different convex sets of probabilities that generate the same preferences under the "unanimity" rule. Figure 11 shows the supporting lines for a set $\mathcal{P}_1$ which arises merely by specifying odds at which betting "on" and "against" the (atomic) events s_j become favorable. The set $\mathcal{P}_1$ is the largest one agreeing with these upper and lower probabilities. As noted by Levi (1980, p. 198), typically, infinitely many convex subsets of $\mathcal{P}_1$ carry the same probability intervals.

Figure 12 illustrates the supporting lines for a set $\mathcal{P}_2$ given, in addition, by bounds on the favorable "called-off" bets $\mathcal{F}_{s_j}c$. $\mathcal{P}_2$ is properly included within $\mathcal{P}_1$, has the same upper and lower probabilities, and is the largest set agreeing with all six pairs of unconditional and conditional odds. As Levi (1974, and 1980, p. 202) points out, we can distin-

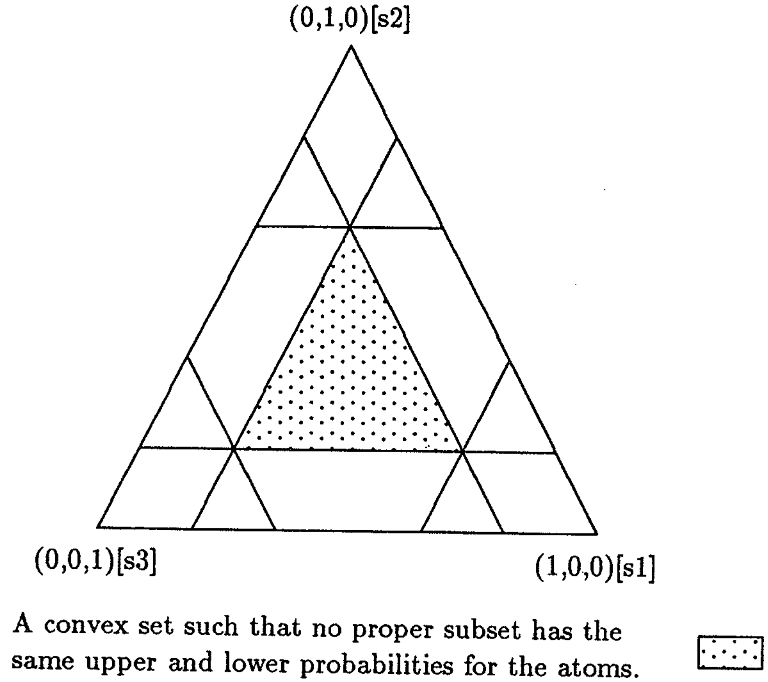


Figure 13. Supporting lines which overdetermine the vertices

guish between two sets having different supporting lines, e.g., $\mathcal{P}_1$ and $\mathcal{P}_2$, with a gamble that is favorable for only one of them.

Figure 13 illustrates how just a few supporting lines can overdetermine the vertices (and thereby all) of a convex set. The simplest case is when the supporting lines corresponding to the upper and lower unconditional odds fix the convex set, $\mathcal{P}_3$, uniquely. That is, there is no proper subset of $\mathcal{P}_3$ with the same upper and lower probabilities. Hence, the set of favorable gambles, $\mathcal{F}$, is fully determined once these upper and lower betting odds are given. (This corrects a minor error in Levi's (1980, p. 202) presentation. There, the set "B_t" has upper and lower unconditional and conditional odds which overdetermine its vertices. Thus, the proper subset "B_t'" does not have the same range of unconditional and conditional odds as "B_t".) We plan to investigate the computational issues relating to the measurement of a convex set, $\mathcal{P}$, using the set of favorable gambles, $\mathcal{F}$. How efficiently can we locate supporting lines which overdetermine the vertices of a set?

V. SUMMARY

We have illustrated a variety of axiomatic and consistency arguments used to justify the normative status of expected utility theory – section I. When (only) the "independence" axiom is denied, inconsistency in sequential choice results – section II. We argue, instead, for a generalization of expected utility theory by relaxing the "ordering" postulate. The resulting theory admits representations in terms of sets of probabilities and utilities – section III. By analogy with the Dutch Book betting argument, we prove that coherence of a partially ordered (strict) preference over gambles (as identified by the set of its strictly "favorable" gambles) is represented by a convex set of probabilities – section IV. Sometimes this representation is fixed by a very few number of comparisons, making measurement feasible.

APPENDIX: REPRESENTATION OF A STRICT PARTIAL ORDER BY A CONVEX SET OF LEXICOGRAPHIC UTILITIES

Defs. Let REW be a set of rewards, REW = $\{r_\alpha : \alpha \leq \beta\}$. A *lottery*, L, is a discrete probability distribution over REW, $L = \{p(\cdot) : p(r_\alpha) \geq 0, \Sigma p(r_\alpha) = 1\}$. Let $Supp(L)$ be the support of $p(\cdot)$. (A *simple lottery* is a lottery with finite support.) Denote by L_{REW} the set of simple lotteries over REW. Given two lotteries $L_1 = \{p_1(\cdot)\}$ and $L_2 = \{p_2(\cdot)\}$, define their convex combination by $L_3 = xL_1 + (1-x)L_2 = \{xp_1(\cdot) + (1-x)p_2(\cdot)\}$. Then, L_{REW} is a (Herstein & Milnor, 1953) *mixture set.*

The following two are our axioms for a strict partial order, $\triangleright$, over L_{REW}.

Axiom 1. $\triangleright$ is a transitive and irreflexive relation on $L_{REW} \times L_{REW}$.

Axiom 2 (Independence). For all L_1, L_2 and L_3, and for each $0 < x \leq 1$:

$$xL_1(1-x)L_3 \triangleright xL_2 + (1-x)L_3 \text{ iff } L_1 \triangleright L_2.$$

Def. When neither $L_1 \triangleright L_2$ nor $L_2 \triangleright L_1$, we say the two lotteries are *incomparable* (by preference), which we denote by $L_1 \sim L_2$.

Incomparability is not transitive, unless $\triangleright$ is a weak order.

Theorem 1. Let REW be a reward set of arbitrary cardinality and let L_{REW} be the set of simple lotteries over these rewards. Let $\triangleright$ be a strict partial order defined over elements of L_{REW}. Then there is an extension of $\triangleright$ to $\triangleright^*$ which is a total ordering of L_{REW} satisfying axiom 2.

Combining Theorem 1 with Hausner's (1954) important result (since a total order is a "pure" weak order), we arrive at the following consequence.

Corollary 1. There is a lexicographic real-valued utility, $\mho$, which agrees with $\triangleright$, i.e., if $L_1 \triangleright L_2$ then $E_\mho[L_1] < E_\mho[L_2]$.

(Note: A lexicographic utility $\mho$ is a (well-ordered) sequence of real-valued utilities, $\mho = \{U_\alpha : U_\alpha$ is a real-valued utility, for each $\alpha < \beta\}$. When $\mho$ is a lexicographic utility, then $E_\mho[L_1] < E_\mho[L_2]$ is said to obtain if $E_{U_\alpha}[L_1] < E_{U_\alpha}[L_2]$ at the first utility U_α in the sequence $\mho$ which gives L_1 and L_2 different expected values, provided one such U_α exists.)

Proof of Theorem 1. Let $\{L_\gamma : \gamma < k$ (γ ranging over ordinals, k a cardinal)$\}$ be a well ordering of L_{REW}. Let $\triangleright$ be a partial order on L_{REW} satisfying axioms 1 and 2. By induction, we define a sequence of extensions of $\triangleright$, $\{\triangleright_\lambda : \lambda \leq k\}$, where each $\triangleright_\lambda$ preserves both axioms and where $\triangleright_k$ is a total order on L_{REW}. The partial order $\triangleright_\lambda$, corresponding to stage λ in the k sequence of extensions, is obtained by contrasting lotteries L_α and L_β, where $\Gamma(\alpha,\beta) = \lambda$ under the canonical well ordering Γ of $k \times k \to k$. We define extensions for successor and limit ordinals separately.

Successor Ordinals. Suppose $\triangleright_\lambda$ satisfies axioms 1 and 2. Let $\Gamma(\alpha,\beta) = \lambda + 1$ and (for convenience) suppose $\max[\alpha,\beta] = \beta$. Define $\triangleright_{\lambda+1}$ as follows.

Case 1: If $\alpha = \beta$, then $\triangleright_{\lambda+1} = \triangleright_\lambda$.

Otherwise,

Case 2: $L_\mu \triangleright_{\lambda+1} L_\nu$ iff either

(i) $L_\mu \triangleright_\lambda L_\nu$ (so $\triangleright_{\lambda+1}$ extends $\triangleright_\lambda$), or

(ii) $L_\alpha \sim_\lambda L_\beta$ & $\exists(0 < x < 1)$ with $xL_\mu + (1-x)L_\beta \triangleright_\lambda$ (or =) $xL_\nu + (1-x)L_\alpha$.

Limit Ordinals. Let $\Gamma(\alpha,\beta) = \lambda < k$, a limit, and (for convenience) again assume $\max[\alpha,\beta] = \beta$.

Case 1: If $\alpha = \beta$, then take $\triangleright_\lambda = \cup_{\delta<\lambda}(\triangleright_\delta)$. That is, $L_\mu \triangleright_\lambda L_\nu$ obtains just in case $\exists(\delta < \lambda) L_\mu \triangleright_\delta L_\nu$.

Case 2: If $\alpha \neq \beta$, then define $\triangleright_\lambda$ as: $L_\mu \triangleright_\lambda L_\nu$ **iff** either (i) $\exists(\delta < \lambda) L_\mu$

$\rhd_\delta L_\nu$ (so $\rhd_\lambda$ extends all preceding $\rhd_\delta$), or (ii) $\forall(\delta < \lambda)L_\alpha \sim_\delta L_\beta$ & $\exists(\delta < \lambda)\exists(0 < x < 1)$ with $xL_\mu + (1 - x)L_\beta \rhd_\delta$ (or =) $xL_\nu + (1 - x)L_\alpha$. Next, we show (by transfinite induction) that $\rhd_\lambda$ satisfies the two axioms, assuming $\rhd(= \rhd_0)$ does. First, consider successor stages where the extension is of the form $\rhd_{\lambda+1}$.

Axiom 1 – irreflexivity. We argue indirectly. Assume, for some lottery L_μ, $L_\mu \rhd_{\lambda+1} L_\mu$. Since $L_\mu \rhd_\lambda L_\mu$ is precluded, by hypothesis of induction, it must be that (ii): $\exists(0 < x < 1)$ with

$$xL_\mu + (1-x)L_\beta \rhd_\lambda (\text{or} =)\ xL_\mu + (1-x)L_\alpha.$$

Since $\rhd_\lambda$ satisfies axiom 2, $L_\beta \rhd_\lambda$ (or =) L_α. If either $L_\beta \rhd_\lambda L_\alpha$ or $L_\beta = L_\alpha$, then $\rhd_{\lambda+1} = \rhd_\lambda$, contradicting the hypothesis $L_\mu \rhd_{\lambda+1} L_\mu$.

Axiom 1 – transitivity. Assume $L_\mu \rhd_{\lambda+1} L_\nu$ and $L_\nu \rhd_{\lambda+1} L_\psi$. There are four cases to consider, since each $\rhd_{\lambda+1}$ relation may obtain in one of two ways. The combination where clause (ii) is used for both provides the greatest generality (the other cases being analyzed in the same way). Thus, we have: $\exists(0 < x, y < 1)$ with

$$xL_\mu + (1-x)L_\beta \rhd_\lambda (\text{or} =)\ xL_\nu + (1-x)L_\alpha$$

and also

$$yL_\nu + (1-y)L_\beta \rhd_\lambda (\text{or} =)\ yL_\psi + (1-y)L_\alpha.$$

Since $\rhd_\lambda$ satisfies axioms 1 and 2, we may "mix" these to yield

$$w(xL_\mu + (1-x)L_\beta) + (1-w)(yL_\nu + (1-y)L_\beta)$$

$\rhd_\lambda$ (or =)

$$w(xL_\nu + (1-x)L_\beta) + (1-w)(yL_\psi + (1-y)L_\beta).$$

Choose $w \cdot x = (1 - w)y$, cancel the common "L_ν" terms (according to axiom 2), regroup (by "reduction") to arrive at: $\exists(0 < v < 1)$

$$vL_\mu + (1-v)L_\beta \rhd_\lambda (\text{or} =)\ vL_\psi + (1-v)L_\alpha,$$

where $v = wx/(1 - y + wy)$. Hence, $L_\mu \rhd_{\lambda+1} L_\psi$, as desired.

Axiom 2. We are to show $L_\mu \rhd_{\lambda+1} L_\nu$ **iff**

$$xL_\mu + (1-x)L_\psi \rhd_{\lambda+1} xL_\nu + (1-x)L_\psi.$$

There are two cases.

Case 1: $L_\mu \rhd_\lambda L_\nu$ occurs just in case $xL_\mu + (1 - x)L_\psi \rhd_\lambda xL_\nu + (1 - x)L_\psi$ (by axiom 2). By the definition of $\rhd_{\lambda+1}$, we obtain the desired result:

$$xL_\mu + (1-x)L_\psi \rhd_{\lambda+1} xL_\nu + (1-x)L_\psi.$$

Case 2: $vL_\mu + (1 - v)L_\beta \rhd_\lambda$ (or =) $vL_\psi + (1 - v)L_\alpha$ occurs just in case

$$yL_\psi + (1-y)(vL_\mu + (1-v)L_\beta) \rhd_\lambda (or =)$$
$$yL_\psi + (1-y)(vL_\psi + (1-v)L_\alpha),$$

according to axiom 2. Choose $y = v(1 - x)/[v(1 - x) + x]$, regroup terms to yield: $w(xL_\mu + (1 - x)L_\psi) + (1 - x)L_\beta \rhd_\lambda$ (or =) $w(xL_\nu + (1 - x)L_\psi) + (1 - x)L_\alpha$, where $w = v/[v(1 - x) + x]$. By the definition of $\rhd_{\lambda+1}$, we obtain the desired result:

$$xL_\mu + (1-x)L_\psi \rhd_{\lambda+1} xL_\nu + (1-x)L_\psi.$$

This establishes axioms 1 and 2 for successor stages, $\rhd_{\lambda+1}$.

The argument with limit stages is similar.

Axiom 1 – irreflexivity. Again, we argue indirectly. Assume $L_\mu \rhd_\lambda L_\mu$. By hypothesis of induction $\neg\exists(\delta < \lambda)\ L_\mu \rhd_\delta L_\mu$. So we may assume $L_\alpha \neq L_\beta$ and $\forall(\delta < \lambda)L_\alpha \sim_\delta L_\beta$ and $\exists(\delta < \lambda)\exists(0 < x < 1)$ with $xL_\mu + (1 - x)L_\beta \rhd_\delta$ (or =) $xL_\mu + (1 - x)L_\alpha$. But by the hypothesis of induction $\rhd_\delta$ satisfies axiom 2, hence, $L_\beta \rhd_\delta$ (or =) L_α, a contradiction.

Axiom 1 – transitivity. Assume $L_\mu \rhd_\lambda L_\nu$ and $L_\nu \rhd_\lambda L_\psi$. Again there are four cases, and again we discuss the most general case where clause (ii) is used to obtain these $\rhd_\lambda$ – preferences. Thus, we have: $\exists(0 < x, y < 1)\ \exists(\delta, \delta' < \lambda)$ with

$$xL_\mu + (1-x)L_\beta \rhd_\delta (\text{or} =)\ xL_\nu + (1-x)L_\alpha$$

and also

$$yL_\nu + (1-y)L_\beta \rhd_{\delta'} (\text{or} =)\ yL_\psi + (1-y)L_\alpha.$$

Without loss of generality, let $\delta = max[\delta, \delta']$. Then

$$yL_\nu + (1-y)L_\beta \rhd_\delta (\text{or} =)\ yL_\psi + (1-y)L_\alpha,$$

since $\rhd_\delta$ extends $\rhd_{\delta'}$. Now, repeat the "mixing" and "cancellation" steps used with the parallel case for successor stages. This yields the desired conclusion: $L_\mu \rhd_\lambda L_\psi$.

Axiom 2. For this axiom, the reasoning is the same as used with axiom 2 in the successor case, modified to apply to the appropriate (preceding) stage $\rhd_\delta$.

Last, define $\rhd_k = \cup_{\delta<k} (\rhd_\delta)$. Hence, $\rhd_k$ is a total order of L_{REW} which satisfies axiom 2. Every two (distinct) lotteries are compared under $\rhd_k$, i.e., $\forall(L_\alpha \neq L_\beta \in L_{REW})\ L_\alpha \rhd_k L_\beta$ or $L_\beta \rhd_k L_\alpha$. □

Next, we state without proof a simple lemma.

Lemma 1. If lexicographic utilities $\mho_1$ and $\mho_2$ both agree with the strict partial order $\rhd$, then so too does their convex mixture $x\mho_1 + (1 - x)\mho_2$. Also, sets of lexicographic utilities generate a strict partial order according to the "unanimity" rule, as we now show.

Lemma 2. Each set of lexicographic utilities $\mathcal{U} = \{\mho: \mho$ is a lexicographic utility over REW$\}$ induces a strict partial order $\rhd_{\mathcal{U}}$ (satisfying axioms 1 and 2) under the "unanimity" rule:

$$L_\alpha \rhd_{\mathcal{U}} L_\beta \textbf{ iff } \forall(\mho \in \mathcal{U})\, E_\mho[L_\alpha] < E_\mho[L - \beta]$$

Proof. The lemma is evident from the fact that each lexicographic utility induces a weak-ordering $\preceq_\mho$ of L_{REW}, satisfying axiom 2, according to the definition:

$$L_\alpha \prec_\mho L_\beta \textbf{ iff } E_\mho[L_\alpha] < E_\mho[L_\beta].$$

Recall, $E_\mho[L_\alpha] < E_\mho[L_\beta]$ obtains if $E_U[L_\alpha] < E_u\,[L_\beta]$ for the first utility U (if one exists) in the sequence $\mho$ which assigns L_α and L_β different expected utilities. Each utility U (hence, $\preceq_U$), supports axiom 2 as:

$$E_U[L_\alpha] < E_U[L_\beta] \textbf{ iff } E_U[xL_\alpha + (1-x)L_\gamma] < E_U[xL_\beta + (1-x)L_\gamma]. \quad \square$$

As is evident from the proof of Theorem 1, if $L \sim L'$, i.e., if neither $L \rhd L'$ nor $L' \rhd L$, then there are alternative extensions of $\rhd$ in which $L \rhd_\delta L'$ and in which $L' \rhd_\delta L$. This observation, together with the two lemmas and Corollary 1, establishes the following representation for strict partial orders $\rhd$.

Theorem 2. Each strict partial order $\rhd$ over a set L_{REW} is identified by a maximal, convex set $\mathcal{U}$ of lexicographic utilities that agree with it. In symbols, $\rhd = \rhd_{\mathcal{U}}$, where $\rhd_{\mathcal{U}}$ is the strict partial order induced by $\mathcal{U}$ under the "unanimity" rule.

Of course, in light of problem (2) (p. 54), it can be that there is a proper (convex) subset $\mathcal{U}' \subset \mathcal{U}$ where $\rhd = \rhd_{\mathcal{U}'}$ as well; hence, the maximality of $\mathcal{U}$ is necessary for uniqueness of the representation.

REFERENCES

Allais, M. (1979) "The So-Called Allais Paradox and Rational Decisions Under Uncertainty," in Allais and Hagen (eds.), *Expected Utility Hypotheses and the Allais Paradox*. D. Reidel: Dordrecht.

Anscombe, F. J., and Aumann, R. J. (1963) "A Definition of Subjective Probability," *Annals of Math. Stat.*, **34**, 199–205.
Aumann, R. J. (1962) "Utility Theory Without the Completeness Axiom," *Econometrica*, **30**, 445–462.
Aumann, R. J. (1964) "Utility Theory Without the Completeness Axiom: A Correction," *Econometrica*, **32**, 210–212.
Bell, D., and Raiffa, H. (1979) "Decision Regret: A Component of Risk Aversion," MS., Harvard University.
Chew Soo Hong and MacCrimmon, K. R. (1979) "Alpha-Nu Choice Theory: A Generalization of Expected Utility Theory," working paper, University of British Columbia.
Cox, D. R. (1958) "Some Problems Connected with Statistical Inference," *Annals of Math. Stat.*, **29**, 357–363.
deFinetti, B. (1937) "La prévision: ses lois logiques, ses sources subjectives," *Annals de l'Institut Henri Poincaré*, **7**, 1–68.
Fishburn, P. C. (1970) *Utility Theory for Decision Making*. Kriefer Publishing Co.: N.Y.
Fishburn, P. C. (1981) "An Axiomatic Characterization of Skew-Symmetric Bilinear Functionals, with Applications to Utility Theory," *Economic Letters*, **8**, 311–313.
Fishburn, P. C. (1983) "Nontransitive Measurable Utility," *J. Math. Psych.*, **26**, 31–67.
Good, I. J. (1952) "Rational Decisions," *J. Royal Stat. Soc. B*, **14**, 107–114.
Hausner, M. (1954) "Multidimensional Utilities," in R. M. Thrall, C. H. Coombs, and R. L. Davis (eds.), *Decision Processes*. Wiley: N.Y.
Herstein, I. N. and Milnor, J. (1953) "An Axiomatic Approach to Measurable Utility," *Econometrica*, **21**, 291–297.
Jeffreys, H. (1971) *Theory of Probability*, 3rd ed. Oxford University Press: Oxford.
Kadane, J., and Sedransk, N. (1980) "Toward a More Ethical Clinical Trial," in Bernardo et al. (eds.), *Bayesian Statistics*. University Press: Valencia.
Kadane, J. B. (ed.) (1996) *Bayesian Methods and Ethics in a Clinical Trial Design*. Wiley: N.Y.
Kahneman, D., and Tversky, A. (1979) "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, **47**, 263–291.
Kannai, Y. (1963) "Existence of a Utility in Infinite Dimensional Partially Ordered Spaces," *Israel J. of Math.*, **1**, 229–234.
Klee, V. L. (1955) "Separation Properties of Convex Cones," *Proc. Amer. Math. Soc.*, **6**, 313–318.
Levi, I. (1974) "On Indeterminate Probabilities," *J. Phil.*, **71**, 391–418.
Levi, I. (1980) *The Enterprise of Knowledge*. MIT Press: Cambridge.
Levi, I. (1986) "The Paradoxes of Allais and Ellsberg," *Economics and Philosophy*, **2**, 23–53.
Lindley, D. V. (1972) *Bayesian Statistics: A Review*. SIAM: Philadelphia.
Loomes, G., and Sugden, R. (1982) "Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty," *Economic J.*, **92**, 805–824.
McClennen, E. F. (1983) "Sure Thing Doubts," in B. Stigum and F. Wenstop

(eds.), *Foundations of Utility and Risk Theory with Applications*. D. Reidel: Dordrecht.
Machina, M. (1982) "'Expected Utility' Analysis Without the Independence Axiom," *Econometrica*, **50**, 277–323.
Machina, M. (1983) "The Economic Theory of Individual Behavior Toward Risk: Theory, Evidence and New Directions," Dept. of Economics, U.C.S.D.: San Diego, CA 92093. Tech. Report #433.
Ramsey, F. P. (1931) "Truth and Probability," in *The Foundations of Mathematics and Other Essays*. Kegan, Paul, Trench, Trubner, and Co. Ltd.: London.
Samuelson, P. (1950) "Probability and the Attempts to Measure Utility," *Economic Review*, **1**, 167–173.
Savage, L. J. (1954) *The Foundations of Statistics*. Wiley: N.Y.
Schick, F. (1984) *Having Reasons*. Princeton Univ. Press: Princeton.
Seidenfeld, T. (1988) "Decision Theory Without Independence or Without Ordering, What Is the Difference?" with discussion, *Economics and Philosophy*, **4**, 267–315.
Sen, A. K. (1977) "Social Choice Theory: A Re-examination," *Econometrica*, **45**, 53–89.
Shimony, A. (1955) "Coherence and the Axioms of Probability," *J. Symbolic Logic*, **20**, 1–28.
Smith, C. A. B. (1961) "Consistency in Statistical Inference and Decision," *J. Royal Stat. Soc. B*, **23**, 1–25.
Suppes, P. (1974) "The Measurement of Belief," *J. Royal Stat. Soc. B*, **36**, 160–175.
Szpilrajn, E. (1930) "Sur l'extension de l'ordre partiel," *Fundamenta Mathematicae*, **16**, 386–389.
von Neumann, J., and Morgenstern, O. (1947) *Theory of Games and Economic Behavior*, 2nd ed. Princeton Univ. Press: Princeton.
Walley, P., and Fine, T. (1979) "Varieties of Modal (Classificatory) and Comparative Probability," *Synthese*, **41**, 321–374.
Wolfenson, M., and Fine, T. (1982) "Bayes-like Decision Making with Upper and Lower Probabilities," *J. Amer. Stat. Assoc.* **77**, 80–88.

1.3

A Representation of Partially Ordered Preferences

TEDDY SEIDENFELD, MARK J. SCHERVISH, AND JOSEPH B. KADANE

ABSTRACT

This chapter considers decision-theoretic foundations for robust Bayesian statistics. We modify the approach of Ramsey, deFinetti, Savage and Anscombe, and Aumann in giving axioms for a theory of *robust* preferences. We establish that preferences which satisfy axioms for robust preferences can be represented by a set of expected utilities. In the presence of two axioms relating to state-independent utility, robust preferences are represented by a *set* of probability/utility pairs, where the utilities are almost state-independent (in a sense which we make precise). Our goal is to focus on preference alone and to extract whatever probability and/or utility information is contained in the preference relation when that is merely a partial order. This is in contrast with the usual approach to Bayesian robustness that begins with a class of "priors" or "likelihoods," and a single loss function, in order to derive preferences from these probability/utility assumptions.

The research of Teddy Seidenfeld was supported by the Buhl Foundation and NSF Grant SES-92-08942. Mark Schervish's research on this project was supported through ONR Contract N00014-91-J-1024 and NSF Grant DMS-88-05676. Joseph Kadane's research was supported by NSF grants SES-89-0002 and DMS-90-05858, and ONR contract N00014-80-J-1851.

Reprinted from *The Annals of Statistics*, 23, no. 6 (1995): 2168–2217, with permission of the Institute of Mathematical Statistics.

The authors thank Peter Fishburn for his careful reading of an earlier version of this essay and for providing us with insightful and extensive feedback. Also, we thank Robert Nau for stimulating exchanges relating to matters of state-dependent utility. We are exceedingly grateful to two *Annals of Statistics* referees, whose patient and detailed readings of our long manuscript resulted in many helpful ideas and useful suggestions.