Journal of Risk and Uncertainty, 11: 17–49 (1995)

A Theory of Coarse Utility

LIPING LIU
School of Business, University of Kansas, Lawrence, Kansas 66045-2003

PRAKASH P. SHENOY
School of Business, University of Kansas, Lawrence, Kansas 66045-2003

Abstract

This article presents a descriptive theory for complex choice problems. In line with the bounded rationality assumption, we hypothesize that decision makers modify a complex choice into some coarse approximations, each of which is a binary lottery. We define the value of a best coarse approximation to be the utility of the choice. Using this paradigm, we axiomatize and justify a new utility function called the *coarse utility function*. We show that the coarse utility function approximates the rank- and sign-dependent utility function. It satisfies dominance but admits violations of independence. It reduces judgmental load and allows flexible judgmental information. It accommodates phenomena associated with probability distortions and provides a better resolution to the St. Petersburg paradox than the expected and rank-dependent theories.

Key words: decision analysis, utility theory, coarse utility function, rank-dependent utilities

Complexity lies deep in the nature of things. Discovering tolerable approximation procedures and heuristics lies at the heart of human intelligence (Simon, 1978). This article presents a descriptive theory of complex decision problems. In line with the bounded rationality assumption (Simon, 1955), we hypothesize that decision makers (DMs) view complex choices approximately. We argue that DMs may not attend to all the outcomes of a many-outcome lottery. Instead, they modify the lottery into some coarse approximations, each of which is a binary lottery with only one nonzero outcome. Such a coarse approximation is shown to be a step distribution function that approximates the original distribution of the lottery. We justify the approximation scheme in terms of complexity, as measured by Shannon's entropy: A many-outcome lottery has more entropy than any of its approximations. We speculate that the more complex a lottery is, the harder it is to assess its utility. Therefore, it is descriptive to evaluate complex lotteries based on their approximations. We also justify the approximation scheme from a prescriptive viewpoint. Approximation is not only an accurate portrayal of much choice behavior, it is also a prescriptively sensible adjustment to the costs and character of information gathering and processing. Relative to exact theories, the coarse utility theory requires not only fewer but also more flexible assessments.

The main purpose of this article is to axiomatize and justify a new utility function called the *coarse utility function*. We do this by means of a mixed rational and plausible approach. We first develop a function that measures the utility of coarse approximations based on some rational and plausible axioms. We then use a set of coarse approximations to approximate a lottery. The value of a best approximation is defined as the utility of the lottery. Let Θ denote the outcome space, which includes $\{0\}$ as the status quo, Θ^+ as the gain subspace, and Θ^- as the loss subspace, i.e., $\Theta = \Theta^- \cup \{0\} \cup \Theta^+$. As in prospect theory (Kahneman and Tversky, 1979), we assume the existence of a value function $v(o)$ over gains and losses with $v(0) = 0$, $v(o) > 0$ for $o \in \Theta^+$, and $v(o) < 0$ for $o \in \Theta^-$. As in the rank- and sign-dependent theory of Luce and Fishburn (1991, 1994) and Tversky and Kahneman (1992), we use different decision weights $W^+(p)$ for gains, and $W^-(p)$ for losses. Given a lottery $L = \{(o, P(o)) \mid o \in \Theta\}$, for any gain $o \in \Theta^+$, let $P(\tilde{o})$ denote the probability of winning o or more. Similarly, for any $o \in \Theta^-$, Let $P(\underset{\sim}{o})$ denote the probability of winning o or less. Then the coarse utility function is represented as follows:

$$U^C(L) = \underset{o \in \Theta^+}{MAX}\, v(o)W^+(P(\tilde{o})) + \underset{o \in \Theta^-}{MIN}\, v(o)W^-(P(\underset{\sim}{o})). \tag{1}$$

The coarse utility function is shown to be very descriptive. It accommodates violations of independence as evidenced by Allais's paradox. It provides a better resolution to the St. Petersburg paradox than the expected and rank-dependent utility theories. The coarse utility function is an approximation of the rank- and sign-dependent utility function of Luce and Fishburn (1991, 1994) and Tversky and Kahneman (1992). When a lottery is a regular prospect (Kahneman and Tversky, 1979), or when a Simon's satisficing value function is assumed, the coarse utility theory and the rank- and sign-dependent utility theory are reconciled. Thus, much of the evidence that supports the rank- and sign-dependent theory also supports the coarse utility theory.

The coarse utility function has been established in a mixed rational and plausible approach. Despite its descriptive focus, it has some normative appeal. The model is shown to underlie a continuous and weak preference relation. It satisfies the requirement of consistency with the first-order stochastic dominance principle. But the underlying preference relation is independent of irrelevant alternatives only when the preference relation is implied by stochastic dominance.

The coarse utility function has some prescriptive advantages. Liu (1994) applied the coarse utility theory to the portfolio selection problem and developed a coarse-utility-based portfolio selection model. The model extends the mean-variance model, in the sense that it imposes no restrictions on asset distributions and preference structures. Under the assumption of normality, it reduces to the mean-variance model. Thus the foundation of modern portfolio analysis (Tobin, 1959; Markowitz, 1959) can also be built on the theory of coarse utility. If all lotteries are normal, then they can be ranked by their means and variances according to the coarse utility criterion. This implication renders the coarse utility theory empirically testable.

An outline of the remainder of this article is as follows. In section 1, we describe and justify the coarse approximation scheme. In section 2, we axiomatize the coarse utility function based on both rational and plausible axioms. In section 3, we describe some basic mathematical properties of the coarse utility function. In section 4, we present empirical evidence in support of the descriptiveness of the coarse utility function. In section 5, we summarize our findings. Finally, in section 6, we give proofs of all results in the article.

1. Coarse approximations

The constructive nature of preferences suggests that preferences are not necessarily procedural-invariant and context-independent. Preferences may be different if lotteries are framed and edited in different ways. The strategies for framing and editing vary as a function of response modes and context complexity (Payne, 1982). In past decades, researchers have actively pursued the fit between decision theories and the DM's information-processing strategies. Guided by observed and/or imagined micro-phenomena in the choice process, theories of choice have become much more refined than ever before. However, there exist gaps between evidence and theory. In general, it remains unclear how individual, task, and context factors affect people's information-processing strategies. These factors still poorly fit generalized utility theories (Camerer, 1992). Specifically, generalized theories often rely on empirical evidence which is accumulated from experiments with simple gambles that abstract away task complexities and their effect on the use of information-processing strategies. In the past, most experiments were conducted on simple gambles such as binary or ternary choices. Generalized theories, on the other hand, were presumably applicable to choices of any complexity. They essentially assume that the information-processing strategies used are invariant across complexities. Unfortunately, evidence does not support this invariance. In simple choice environments, various simple heuristics may be available that closely approximate the responses from a utility theory. For example, anchoring and adjustment models can closely approximate expected utility theory in some cases (Hershey and Schoemaker, 1985; Johnson and Payne, 1985; Goldstein and Einhorn, 1987; Johnson and Schkade, 1989). However, when choice problems are complex, such as selecting portfolios or deductibles in insurance policies, the calculations and problems structures will just become too complex for optimal solutions (Schoemaker, 1993).

Complexity raises an obvious inconsistency between our methodologies of formulating utility theories and the bounded rationality assumption. Traditionally, a utility function is formulated by implicitly assuming that DMs are able to attend to all the outcomes of any complex gambles. They view a lottery as a set of outcome-probability pairs. Each outcome-probability pair contributes to the utility of the lottery, i.e., summing the utility of each outcome-probability pair gives the utility of the lottery. This paradigm of formulation reigned for several centuries, from the proposal of Bernoulli's expected utility to the formulation of early decision weight models (Edwards, 1955, 1962; Fellner, 1961; Handa, 1977; Karmarkar, 1978), and to the more recent formulation of various rank-dependent utility theories (Kahneman and Tversky, 1979; Quiggin, 1982; Yaari, 1987; Luce and Fishburn, 1991, 1994; Tversky and Kahneman, 1992). If we interpret the contribution of each outcome-probability pair more generally to include a weighted value or a regret, then this dominant weighted-average paradigm is also shared by many other nondecision-weight models, such as the weighted linear model of Chew (1983), the expected regret model of Bell (1982) and Loomes and Sugden (1982), and the SSA model of Fishburn (1986).

On the other hand, the bounded rationality assumption implies that people have difficulties in anticipating or considering all options and all information (March, 1978).

They tend to discover approximation procedures and heuristics to simplify decision problems. Simon (1978) argues that attention is sometimes a scarce resource. We cannot afford to attend to information simply because it is there. Approximation is an economical way of using our attention resource. Many behavioral researchers have found that people's internal representations of problems are simple, even though the environment is complex (Brunswik, 1952; Hammond, McClelland, and Mumpower, 1980). For example, Kunreuther et al. (1978) found that learning about flood insurance from friends or neighbors was a key determinant of purchase. Empirical evidence shows that human information processing is seldom more complex than a linear model (Dawes and Corrigan, 1978; Dawes, Faust, and Meehl, 1989). Cohen, Jaffray, and Said (1985) found that people can seldom comprehend probabilities beyond very coarse descriptions such as "very likely," "likely," and "not likely."

At this time, little is known about how people approach complex decision problems. However, our informal observations support the generalization that DMs do not attend to all the outcomes of a many-outcome gamble. For example, some lotteries are often pessimistically valued by their bad outcomes, and others are often optimistically valued by their good outcomes. DMs' intrinsic risk attitudes explain such phenomena (Schoemaker, 1993). We argue that their focused attention also matters. As reported by Tversky, Sattath, and Slovic (1988), choice responses are often dominated by prominent attributes, which tend to form the bases for compelling reasons or arguments for the choice made. In line with this observation, we hypothesize that DMs do not attend to all the outcomes of a many-outcome lottery. Instead, they approximate a many-outcome lottery by some coarse approximations, each of which is a binary gamble with one nonzero outcome. The approximation scheme depends on whether a lottery involves only gains, or only losses, or mixed gains and losses. We describe it case by case in the following.

In the case of gains, consider lottery L_1 with ten mutually exclusive monetary outcomes, \$0, \$100, \$200, . . . , \$900. Assume that each of these outcomes is equally likely. Traditionally, it has been assumed that DMs evaluate L_1 based on the unedited information that L_1 pays \$100 10% of the time, \$200 another 10% of the time . . . , and \$900 still another 10% of the time. That is, L_1 is the set of outcome-probability pairs as follows: $L_1 = \{(\$0, 0.1), (\$100, 0.1), (\$200, 0.1), \ldots, (\$900, 0.1)\}$. We assume that DMs evaluate L_1 based on the edited information that L_1 will pay at least \$100 90% of the time, at least \$200 80% of the time, . . . , or at least \$900 10% of the time. Similarly, in the case of losses, consider lottery $L_2 = \{(\$0, 0.1), (-\$100, 0.1), (-\$200, 0.1), \ldots, (-\$900, 0.1)\}$. We assume that DMs evaluate L_2 based on the edited information that L_2 will make you lose at least \$100 90% of the time, at least \$200 80% of the time, . . . , or at least \$900 10% of the time. In the case of mixed gains and losses, we assume that DMs first code outcomes as gains and losses in the editing phase. Then a lottery involving mixed gains and losses is decomposed into a joint receipt of a positive part, obtained by replacing losses by the status quo, and a negative part, obtained by replacing gains by the status quo. For example, the mixed lottery $\{(\$0, 0.1), (\$100, 0.5), (-\$200, 0.4)\}$ can be decomposed into $\{(\$0, 0.5), (\$100, 0.5)\}$ and $\{(\$0, 0.6), (-\$200, 0.4)\}$. In summary, we assume that, although a lottery is presented to a DM as a set of outcome-probability pairs, it is represented and processed in cumulative or decumulative terms. This assumption has been supported by many psychological experiments. As Lopes (1984, 1987) has argued, Lorenz curves, which are normalized cumulative distribution curves, can capture cumulative features of risky choices that are salient to people when they judge or choose risks.[1] Some verbal protocols in Schneider and Lopes (1986) reveal that people tend to talk as though they view lotteries in terms of the cumulative likelihood of meeting or exceeding a certain outcome or the decumulative likelihood of getting zero or bad outcomes.

We now formally describe and justify the above approximation scheme. Consider lottery L_1 described above. Consider an outcome of L_1, say \$300, to focus on. Then the lottery L_1 gives a monetary outcome of less than \$300 with probability 0.3, and a monetary outcome of \$300 or more with probability 0.7. Therefore, L_1 is approximated by a binary lottery $\{(<\$300, 0.3), (\geq \$300, 0.7)\}$. In this binary lottery, $(\geq \$300, 0.7)$ is the outcome-probability pair that is of interest to the DM. Relative to $(\geq \$300, 07)$, $(<\$300, 0.3)$ represents a group of moderate and zero payoffs, and its value can be ignored. Hence, we further approximate $(\geq \$300, 0.7)$ by $(\$300, 0.7)$ and $(<\$300, 0.3)$ by $(\$0, 0.3)$. Consequently, we obtain another binary lottery $L_1(\$300) = \{(\$0, 0.3), (\$300, 0.7)\}$. It is clear that the lottery $L_1(\$300)$ is an approximation of the original lottery L_1. We call $L_1(\$300)$ a *coarse approximation* of L_1. Given any lottery, we can usually derive many coarse approximations of it, as above, by focusing on different outcomes. The lottery L_1 can be approximated by ten coarse approximations—one for each outcome of L_1—as follows: $L_1(\$0) = \{(\$0, 1)\}$, $L_1(\$100) = \{(\$0, 0.1), (\$100, 0.9)\}$, $L_1(\$200) = \{(\$0, 0.2), (\$200, 0.8)\}, \ldots, L_1(\$800) = \{(0, 0.8), (\$800, 0.2)\}$, and $L_1(\$900) = \{(\$0, 0.9), (\$900, 0.1)\}$.

By the reduction principle (Fishburn, 1988), a lottery is simply a probability distribution over a set of outcomes. In terms of cumulative probabilities, we can regard a coarse approximation as a step function that approximates the cumulative probability distribution of the original lottery. For example, consider lottery L_1 as described above and its approximation $L_1(\$500) = \{(\$0, 0.5), (\$500, 0.5)\}$. Their cumulative probability distributions are shown by figure 1.

A psychological justification for approximating a lottery lies in the fact that a lottery is more complex than its approximation. If we regard a lottery $L = \{(o_1, p_1), (o_2, p_2), \ldots, (o_n, p_n)\}$ as a random system having the state o_i with probability p_i, then entropy $H(L) = \Sigma\{-p_i \log(p_i) \mid i = 1, 2, \ldots, n\}$ can be used as a standard measure of complexity of the lottery L (Shannon, 1948). It follows from this formula that a lottery has more entropy than its approximation. For example, the entropy of the lottery $L_1 = \{(\$0, 0.1), (\$100, 0.1), (\$200, 0.1), \ldots, (\$900, 0.1)\}$ is 3.3, while the entropy of any coarse approximation is less than or equal to 1. We speculate that the more complex a lottery is, the harder it is to assess its utility. This speculation is supported by the work of Hogarth and Einhorn (1990), in which they demonstrate that the amount of mental simulation used in assessing the utility of a lottery increases with complexity. Consequently, people may be more comfortable in assessing coarse approximations of a complex lottery than the lottery itself.

Lottery $L_1(\$0)$ has only one outcome-probability pair, (\$0, 1). The remaining nine approximations are binary lotteries. Given the nature of each binary approximation, we can simply describe it by one outcome-probability pair. For example, lottery $L_1(\$100) = \{(\$0, 0.1), (\$100, 0.9)\}$ is described completely by (\$100, 0.9), $L_1(\$200)$ by (\$200, 0.8), and so on. Thus the set of ten approximations of L_1 can be described by a set of ten outcome-probability pairs, namely, $\{(\$0, 1), (\$100, 0.9), (\$200, 0.8), \ldots, (\$900, 0.1)\}$. In general, if

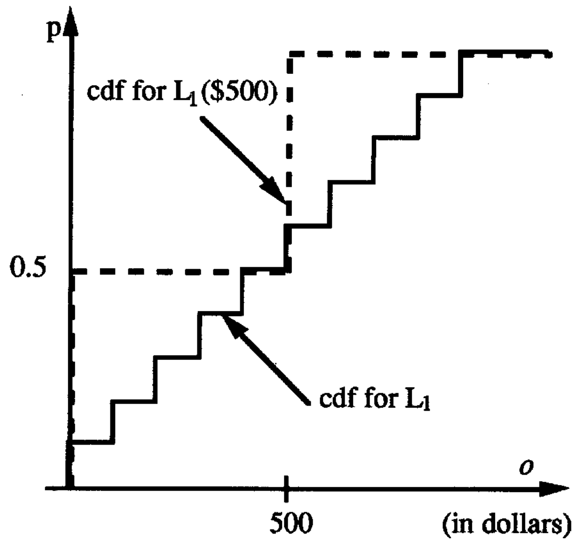

Figure 1. The cumulative distributions of lottery L_1 and its approximation $L_1(\$500)$.

a lottery involves only gains, i.e., $L = \{(o, P(o)) \mid o \in \Theta^+ \cup \{0\}\}$, then $(o, P(\tilde{o}))$ is a coarse approximation for any outcome $o \in \Theta^+ \cup \{0\}$, where $P(\tilde{o})$ is the probability of winning o or more.

We can similarly describe and justify coarse approximations for a lottery involving only losses. Consider lottery L_2 described earlier. Consider an outcome of L_2, say $-\$300$ to focus on. Then the corresponding coarse approximation for L_2 is binary lottery $L_2(-\$300) = \{(-\$300, 0.7), (\$0, 0.3)\}$. L_2 can be also approximated by other coarse approximations—one for each outcome of L_2—as follows: $L_2(\$0) = \{(\$0, 1)\}$, $L_2(-\$100) = \{(-\$100, 0.9), (\$0, 0.1)\}$, $L_2(-\$200) = \{(-\$200, 0.8), (\$0, 0.2)\}$, ... , and $L_2(-\$900) = \{(-\$900, 0.1), (\$0, 0.1)\}$. In terms of decumulative probabilities, each coarse approximation is a step function that approximates the decumulative probability distribution of the original lottery L_2. For example, consider lottery L_2 as described above and its approximation $L_2(-\$300) = \{(-\$300, 0.7), (\$0, 0.3)\}$. Their decumulative probability distributions are shown by figure 2. We can simply describe each binary approximation by one outcome-probability pair. For example, lottery $L_2(-\$100)$ is described completely by $(-\$100, 0.9)$, $L_2(-\$200)$ by $(-\$200, 0.8)$, and so on. Thus, the set of ten approximations of L_2 can be described by a set of ten outcome-probability pairs, namely, $\{(\$0, 1), (-\$100, 0.9), (-\$200, 0.8) \ldots, (-\$900, 0.1)\}$. In general, if a lottery involves only losses, i.e., $L = \{(o, P(o)) \mid o \in \Theta^- \cup \{0\}\}$, then $(o, P(\underset{\sim}{o}))$ is a coarse approximation for any outcome $o \in \Theta^- \cup \{0\}$, where $P(\underset{\sim}{o})$ is the probability of winning o or less.

A complex lottery involving mixed gains and losses is decomposed into the joint receipt of a positive part involving only gains, and a negative part involving only losses, which are then approximated according to the above schemes.

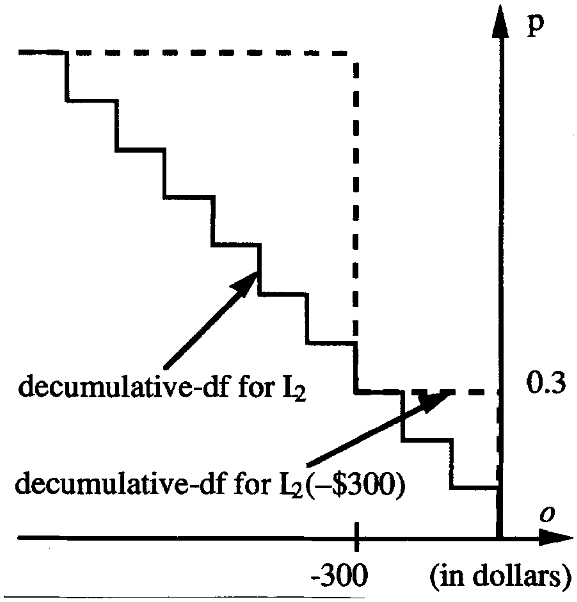

Figure 2. The decumulative distributions of lottery L_2 and its approximation $L_2(-\$300)$.

The coarse approximation scheme can also be justified from its potential prescriptive properties (Bell, Raiffa, and Tversky, 1988). Theories of rational choice stand up on the grounds of two critical assessments: assessments about future consequences of current actions and assessments about future preferences for those consequences (Savage, 1954; Thompson, 1967). Neither assessment is necessarily easy. Both assessments are further complicated when uncertainty is involved. Exact theories usually require the consequences and their probabilities to be assessed with precision. This seems to be too difficult in practice. The number of outcome-probability pairs to be assessed may be infeasibly large, and the consistency of information to be assessed may be beyond the limited rationality of human beings. According to the dominant paradigm for formulating utility theories, each possible outcome-probability pair has a value-added contribution to the utility of a choice. Thus, ignoring or mis-assessing any outcome will miscalculate its utility. To be consistent with the theories, all the possible outcomes and the corresponding probabilities of each choice should be assessed precisely. Otherwise, using Howard's (1992) terminology, no warranties can be obtained. For example, Brockett and Kahane (1992) show that, for any lottery (continuous or discrete, unimodal or asymmetric), even if an approximation matches any number of its moments exactly, it may not approximate its expected utility (see also Liu, 1993; Keefer, 1994). Applying the Markov-Krein Theorem (Karlin and Studden, 1966) to the Choquet integral, one may also show that similar results hold for generalized theories. In contrast, as suggested by its formulation, the coarse utility function requires much fewer assessments. The coarse utility of a lottery is determined only by its best coarse approximation. Therefore, in essence, only one assessment corresponding to a best approximation is critical to the utility calculation. This feature of the coarse utility theory implies the substantial reduction of judgmental load. It also implies the improvement of judgmental information quality by suggesting the focus of deliberation.

2. The coarse utility function

In this section, we define and justify the coarse utility function using a mixed rational and plausible approach. We first establish a utility function for coarse approximations of lotteries involving only gains and losses. We then define the utility of a positive or negative lottery to be the utility of its best coarse approximation, and the utility of a mixed lottery to be the sum of the utilities of its positive and negative decompositions.

Each coarse approximation is an outcome-probability pair (o, p) in $\Theta \times [0, 1]$. If $o \in \Theta^+ \cup \{0\}$, then p is the probability of winning o or more. If $o \in \Theta^- \cup \{0\}$, p is the probability of winning o or less. Let $\succsim$ be a binary preference relation on $\Theta \times [0, 1]$. We define a utility function V for the coarse approximations as a mapping V on $\Theta \times [0, 1]$ such that

$$V(o_1, p_1) \geq V(o_2, p_2) \text{ if and only if } (o_1, p_1) \succsim (o_2, p_2).$$

Such a utility function can be regarded as a bi-attribute value function, where outcomes and likelihoods are the two criteria considered and maximized. Let $\sim$ and $\succ$ be, respectively, the indifference and the strictly preference relations induced from $\succsim$. We assume the five axioms as follows:

Axiom 1. Θ is a connected and separate topological space.

Axiom 2. $\succsim$ is a weak order on $\Theta \times [0, 1]$.

Axiom 3. $\succsim$ is continuous on $\Theta \times [0, 1]$.

Axiom 4. If (o_1, p_1) first-degree dominates (o_2, p_2), then $(o_1, p_1) \succ (o_2, p_2)$.

Axiom 5. If o_1, o_2, and o_3 are nonzero and p_1, p_2, and $p_3 > 0$, $(o_1, p_1) \sim (o_2, p_2)$ and $(o_2, p_3) \sim (o_3, p_1)$ implies $(o_1, p_3) \sim (o_3, p_2)$.

These axioms are intuitively acceptable at a normative level. Axioms 1-3 are the sufficient and necessary conditions for the existence of a continuous and order-preserving function determined up to a strictly increasing transformation (Debreu, 1954). Axiom 4 is intuitive and normative. Axiom 5 is called the *cancellation assumption* (Roskies, 1964) or the *Thomsen condition* (Krantz et al., 1971), which is weaker than the counterpart assumed by Debreu (1960) and Luce and Tukey (1964).

2.1. *The coarse utility of positive lotteries*

If a lottery involves only gains, i.e., $L = \{(o, P(o)) \mid o \in \Theta^+ \cup \{0\}\}$, then $(o, P(\tilde{o}))$ is a coarse approximation for any point $o \in \Theta^+ \cup \{0\}$. Therefore, any coarse approximation of a positive lottery is a point in the space $[\Theta^+ \cup \{0\}] \times [0, 1]$. Theorem 1 shows that there is a multiplicative function V, which is unique up to a transformation V^α where $\alpha > 0$ and can be used to measure the preference on $[\Theta^+ \cup \{0\}] \times [0, 1]$ qualitatively characterized by Axioms 1-5.

Theorem 1. Suppose o^* is a most preferred outcome in Θ^+, and any outcome o in Θ^+ is strictly preferred to the status quo 0. Axioms 1-5 hold iff there exists an order-preserving function $V(o,p)$, which is continuous in $\Theta^+ \times (0, 1]$, such that:

$$V(o,p) = V(o, 1)V(o^*,p), V(o, 0) = 0, V(o^*, 1) = 1, V(0,p) = 0, \tag{2}$$

$\forall (o, p) \in [\Theta^+ \cup \{0\}] \times [0, 1]$. If V' is another function satisfying (2), then $V' = V^\alpha$ where $\alpha > 0$.

Since $V(o, 1)$ measures the value of o under certainty, it is an ordinary value function and denoted by $v(o)$. It is easy to see from (2) that $v(o)$ is a strictly increasing function of o, and $v(0) = 0$ and $v(o^*) = 1$. $V(o^*, p)$ measures the preference value of probability p, with which the most preferred outcome o^* is obtained. $V(o^*, p)$ is a strictly increasing function of probability p and denoted by $W^+(p)$. According to (2), $W^+(0) = 0$ and $W^+(1) = 1$. In these notations, (2) becomes

$$V(o,p) = v(o)W^+(p), v(0) = 0, v(o^*) = 1, W^+(0) = 0, W^+(1) = 1. \tag{3}$$

Therefore, for lottery $L = \{(o, P(o)) \mid o \in \Theta^+ \cup \{0\}\}$, its coarse approximation $(o, P(\tilde{o}))$, corresponding to the outcome $o \in \Theta^+ \cup \{0\}$, has the utility value $V(o, P(\tilde{o})) = v(o)W^+(P(\tilde{o}))$. Since there are many such approximations to L, we define the coarse utility of the lottery L to be the maximum utility value of the coarse approximations:

$$U^C(L) = \underset{o \in \Theta^+}{MAX}\ v(o)W^+(P(\tilde{o})). \tag{4}$$

Note that $v(o)W^+(P(\tilde{o})) = 0$ if $o = 0$. Therefore, $U^C(L)$ is well defined in (4), even though we take maximum over Θ^+ rather than $\Theta^+ \cup \{0\}$. As argued in section 1, we essentially approximate lottery $L = \{(o, P(o)) \mid o \in \Theta^+ \cup \{0\}\}$ by a binary lottery $\{(0, 1 - P(\tilde{o})), (o, P(\tilde{o}))\}$. The approximation is conservative. Then in (4) we assess the utility of a lottery by the maximum utility of its coarse approximations. Choosing the maximum utility in (4) offsets the conservatism.

Equation (4) is a normative definition for the coarse utility of a positive lottery. Formally, to compute a coarse utility value requires one to solve an optimization problem. This is obvious beyond the bounded rationality. However, we see definition (4) as a descriptive approximation of DM's intention to search for better coarse approximations.

As we argued before, a binary lottery, say $L(o) = (o, P(\tilde{o}))$, is only one approximation of the lottery L. There are several such approximations and some are better than others. In figure 3, we plot the cumulative probability distribution of L and the cumulative probability distribution of the approximation $L(o)$. We measure the goodness of approximation $L(o)$ by the total absolute error $E(L(o)) = A_1 + A_2$, where A_1 and A_2 are the areas between the cumulative distribution of L and that of $L(o)$, as shown in figure 3. If we are correct in hypothesizing that a DM uses the coarse approximation scheme to evaluate L, he or she will attempt to obtain a best approximation $L(o)$ (for $o \in \Theta^+ \cup \{0\}$) that has the smallest total absolute error $E(L(o))$. It is clear from figure 3 that the total area between the cumulative distribution curve and the horizontal line $p = 1$, i.e., $A_1 + A_2 + A_3$, is fixed. $E(L(o)) = A_1 + A_2$ is minimized only when $A_3 = oP(\tilde{o})$ is maximized. Because of the subjective perception effect, o is perceived as $v(o)$, and $P(\tilde{o})$ is perceived as $W^+(P(\tilde{o}))$. To choose a best coarse approximation, people maximize $v(o)W^+(P(\tilde{o}))$. The maximum value itself, by (4), is the utility value of the best approximation. Therefore, we justify (4), in the sense that a DM attempts to evaluate a lottery based on approximations which are as close as possible.

2.2. *The coarse utility of negative lotteries*

If a lottery involves only losses, i.e., $L = \{(o, P(o)) \mid o \in \Theta^- \cup \{0\}$, $(o, P(\underset{\sim}{o}))$ is a coarse approximation for any outcome $o \in \Theta^- \cup \{0\}$. Therefore, any coarse approximation of a negative lottery is a point in the space $[\Theta^- \cup \{0\}] \times [0, 1]$. Theorem 2 shows that there exists a multiplicative function V, which is unique up to a transformation $-(-V)^\alpha$ where $\alpha > 0$ and which can be used to measure the preference on $[\Theta^- \cup \{0\}] \times [0, 1]$ qualitatively characterized by Axioms 1–5.

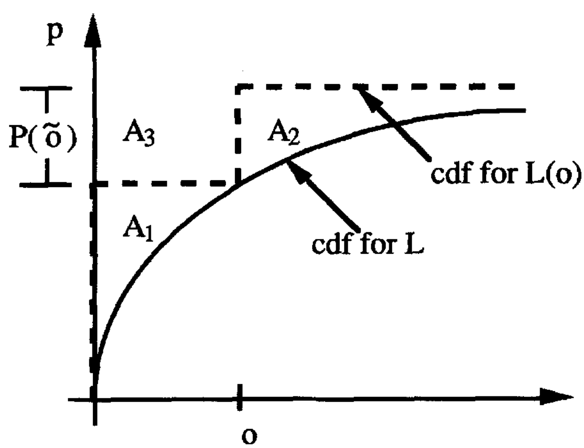

Figure 3. The measure of approximation errors for a positive lottery.

Theorem 2. Suppose o_* is a least preferred outcome in Θ^-, and any outcome o in Θ^- is strictly less preferred than the status quo 0. Axioms 1-5 hold iff there exists an order-preserving function $V(o,p)$, which is continuous in $\Theta^- \times (0, 1]$, such that:

$$V(o,p) = -V(o,1)V(o_*,p), V(o,0) = 0, V(o_*,1) = -1, V(0,p) = 0, \tag{5}$$

$\forall(o, p) \in [\Theta^- \cup \{0\}] \times [0, 1]$. If V' is another function satisfying (5), then $V' = -(-V)^\alpha$ with $\alpha > 0$.

$V(o, 1)$ measures the value of o under certainty. Therefore, $V(o, 1)$ is an ordinary value function of losses and is denoted by $v(o)$. It is easy to see from (5) that $v(o)$ is a strictly increasing function of o, and $v(0) = 0$ and $v(o_*) = -1$. $V(o_*,p)$ measures the preference value of probability p, with which the most undesirable loss o_* is obtained. $V(o_*,p)$ is a strictly decreasing function of probability p. Let $W^-(p) = -V(o_*,p)$. Then $W^-(p)$ can be regarded as the preference value of probability p, with which the most undesirable loss o_* is avoided. $W^-(p)$ is a strictly increasing function, and $W^-(0) = 0$ and $W^-(1) = 1$. In these notations, (5) becomes

$$V(o,p) = v(o)W^-(p), v(0) = 0, v(o_*) = -1, W^-(0) = 0, W^-(1) = 1. \tag{6}$$

Therefore, applying (6) to a coarse approximation of L, say $(o, P(\underset{\sim}{o}))$, yields $V(o, P(\underset{\sim}{o})) = v(o)W^-(P(\underset{\sim}{o}))$. Since there are many such coarse approximations to L, we define the coarse utility of L to be the minimum utility value of the coarse approximations:

$$U^C(L) = \underset{o \in \Theta^-}{MIN} \; v(o)W^-(P(\underset{\sim}{o})). \tag{7}$$

Note that $v(o)W^-(P(\underset{\sim}{o})) = 0$ if $o = 0$. Therefore, $U^C(L)$ is well defined in (7), even though we take minimum only over Θ^- rather than $\Theta^- \cup \{0\}$. As argued in section 1, we approximate lottery $L = \{(o, P(o)) \mid o \in \Theta^- \cup \{0\}\}$ by binary lotteries $\{(0, 1 - P(\underset{\sim}{o})), (o, P(\underset{\sim}{o}))\}$. The approximation overestimates the value of L. Then in (7) we assess the utility of a lottery by the minimum utility of its coarse approximations. Choosing the minimum utility in (7) offsets the overestimation.

We can similarly give a behavioral justification for (7), in the sense that it represents a descriptive approximation of a DM's intention to search for better coarse approximations. But this time we define a better approximation to be the one whose decumulative distribution is closer to that of the original lottery L. Given a negative lottery L and one of its approximations $L(o) = (o, P(\underset{\sim}{o}))$, in figure 4, we plot the decumulative probability distribution of L and that of approximation $L(o)$. We measure the goodness of approximation $L(o)$ by the total absolute error $E(L(o)) = A_1 + A_2$, where A_1 and A_2 are the areas between the decumulative distribution of L and that of $L(o)$, as shown in figure 4. A DM attempts, by our hypothesis, to find a best approximation $L(o)$ (for $o \in \Theta^-$ or $o = 0$) that has the smallest total absolute error $E(L(o))$. It is clear from figure 4 that the total area between the decumulative distribution curve of L and the horizontal line $p = 1$, i.e., $A_1 + A_2 + A_3$, is fixed. $E(L(o)) = A_1 + A_2$ is minimized only when A_3 is maximized. That corresponds to the minimization of $oP(\underset{\sim}{o})$, because o is negative. Due to

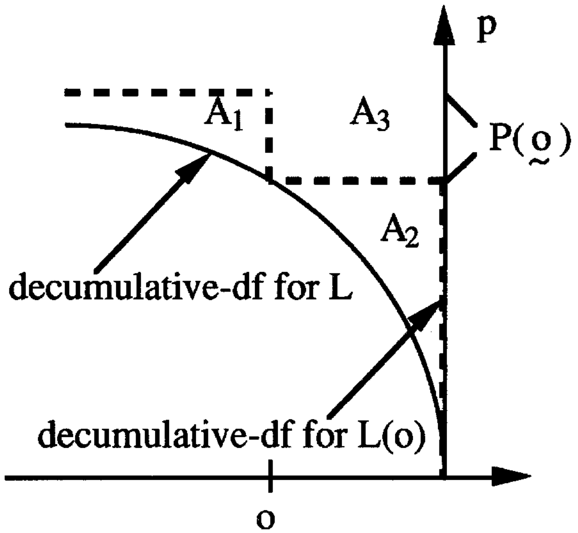

Figure 4. The measure of approximation errors for a negative lottery.

perceptional effect, o is perceived as $v(o)$, and $P(\underset{\sim}{o})$ is perceived as $W^-(P(\underset{\sim}{o}))$. To choose a best coarse approximation, people minimize $v(o)W^-(P(\underset{\sim}{o})$. The minimum value itself, by (7), is the utility value of the best approximation of the lottery L.

2.3. *The coarse utility of mixed lotteries*

Following the decomposition axiom of Luce and Fishburn (1991), we treat a lottery with mixed gains and losses as indifferent to the joint receipt of its gains pitted against the status quo and of its losses against the status quo. Of course, using the double matching and the co-monotonic independence axioms of Tversky and Kahneman (1992) and Wakker and Tversky (1991), we can also establish that the utility of a lottery is the sum of the utilities of both its positive and negative parts.

Suppose $L = \{(o, P(o)) \mid o \in \Theta\}$ is a mixed lottery. Then, its positive part, denoted by L^+, is

$$L^+ = \{(o, P(o)) \mid o \in \Theta^+\} \cup \left\{\left(0, 1 - \sum_{o \in \Theta^+} P(o)\right)\right\},$$

and its negative part, denoted by L^-, is

$$L^- = \{(o, P(o)) \mid o \in \Theta^-\} \cup \left\{\left(0, 1 - \sum_{o \in \Theta^-} P(o)\right)\right\},$$

In light of Tversky and Kahneman (1992) and Luce and Fishburn (1991), we assume that the utility of a mixed lottery is equal to the sum of its decompositions.

Axiom 6. Suppose L^+ and L^- are the positive and the negative duplex decompositions of L, respectively. Then, $U^C(L) = U^C(L^+) + U^C(L^-)$.

Axiom 6 was called *duplex decomposition* by Slovic (1967) and Slovic and Lichtenstein (1968). As Luce and Fishburn (1991) have argued, it is not rational but plausible. Some empirical data support its descriptive accuracy (Slovic, 1967; Slovic and Lichtenstein, 1968; Cho, Luce, and von Winterfeldt, 1993). Notice that $v(0) = 0$. Thus, combining (4), (7), and Axiom 6 yields

$$U^C(L) = \underset{o\in\Theta^+}{MAX}\, v(o)W^+(P(\tilde{o})) + \underset{o\in\Theta^-}{MIN}\, v(o)W^-(P(\underset{\sim}{o})).$$

2.4. *Psychophysics of perception*

According to Luce and Fishburn (1991) and Tversky and Kahneman (1992), a binary lottery $\{(0, 1 - p), (o, p)\}$ with the single nonzero outcome o has rank- and sign-dependent utility $v(o)W^+(p)$ if o is a gain, or $v(o)W^-(p)$ if o is a loss. It follows from (3) and (6) that $V(o^*, p)$ is a weighting function for gains, and $-V(o_*, p)$ is a weighting function for losses, in the spirit of Kahneman and Tversky (1979). Therefore, $W^+(p)$ and $W^-(p)$, which are defined in this article as $W^+(p) = V(o^*,p)$ and $W^-(p) = -V(o_*,p)$, can be called decision weights, consistent with the literature of rank- and sign-dependent utility theories. Also, in line with Kahneman and Tversky (1979), Luce and Fishburn (1991), and Tversky and Kahneman (1992), we assume a value function $v(o)$ over gains and losses, with $v(0) = 0, v(o) > 0$ for gains, and $v(o) < 0$ for losses. We allow different decision weights $W^+(p)$ and $W^-(p)$ for gains and losses, with $W^+(0) = W^-(0) = 0$, and $W^+(1) = W^-(1) = 1$.

The coarse utility function stands up on the ground of the normative and plausible Axioms 1–6. However, to test its descriptive accuracy, we need to study the features of the value function $V(o,p)$ from the psychophysical perspective. Specifically, we need to know how real people perceive the value function $v(o)$ and the weighting functions $W^+(p)$ and $W^-(p)$.

Based on the assumption that equal relative changes in wealth are equally significant, Bernoulli proposed a logarithmic value function that is concave everywhere. Allais empirically justified Bernoulli's proposition. However, early work on the value assessments found a bizarre phenomenon: most value functions assessed are concave for gains and convex for losses. This prevalent finding led Kahneman and Tversky (1979) and Tversky and Kahneman (1992) to formally propose that the value function is S-shaped and is steeper for losses than for gains. They justify this assumption based on the principles of diminishing sensitivity and loss aversion. Of course, as Fishburn (1988) has pointed out, the above finding is by no means universal. Cohen, Jaffray, and Said (1985) found no correlation between attitudes toward gains and losses. Leland (1986, 1988) and Friedman (1989) argue and justify that the S-shaped value function is only an approximation of "true" preferences, given the constraint of a subject's inexperience and/or cognitive limitations.

Extensive experiments have reported that a weighting function takes a skew S-shape, as shown in figure 5 (Kahneman and Tversky, 1979). Notice that the slope of the skew S-curve is initially very large, then becomes smaller, but finally increases as p approaches 1. In other words, the curvature is negative when p is close to zero, and positive when p is close to one. Tversky and Kahneman (1992) assume that decision weights measure the contributions of events to the desirability of prospects. They argue that the principle of diminishing sensitivity also applies to the interpretation of the skew S-shaped pattern of weighting functions. Currently, a large body of evidence seems to support the claim[2] that a normative theory that does not consider probability distortions cannot be an accurate descriptive theory (Fishburn, 1988).

In this article, $v(o) = V(o, 1)$, $W^+(p) = V(o^*, p)$, and $W^-(p) = -V(o_*, p)$, while $V(o, p)$ measures a person's preference of the binary lottery $\{(0, 1 - p), (o, p)\}$. Therefore, observed patterns of values and weights are determined by the pattern of the function $V(o, p)$, which, according to psychophysics, is in turn determined by individual perceptional sensitivity. If one subscribes to the psychophysical perspective, one can apply the principles of diminishing sensitivity and loss aversion to the perception of $V(o, p)$ and provide a unified explanation to the S-shaped pattern of value functions and the skew S-shaped pattern of weighting functions. For example, $W^+(p) = V(o^*, p)$ measures the utility value of receiving lottery $\{(0, 1 - p), (o^*, p)\}$. When p is close to zero, people naturally compare it with zero and value it as a gain of certainty from zero. On the other hand, when p is close to one, people naturally compare it with one and value it as a loss of certainty from one. $W^+(p)$ is concave in the frame of gains of certainty, and convex

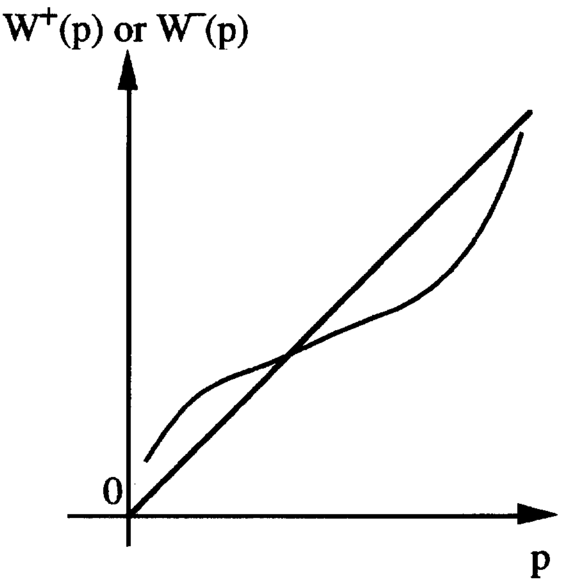

Figure 5. Skew S-curve of weighting functions.

in the frame of losses of certainty. This is consistent with the S-shaped value function for monetary outcomes which Tversky and Kahneman refer to as the reflection effect. Also, it is obvious from $W^+(p) = V(o^*, p)$ and $W^-(p) = -V(o_*, p)$ that the size and the sign of the extreme outcomes o^* and o_* enter into the valuation of probabilities. The behavior of perceiving p will be different given different absolute sizes of o^* and o_*. The difference is attributed to probability $\times$ utility interactions by many theorists and supported by the venture theory of Hogarth and Einhorn (1990) and others.

Note that the representation of values and weights as functions of o and p is by no means complete. According to psychophysics (Sinn, 1983), perceptional sensitivity depends not only on the magnitude of the stimulus o and p, but also on some important personal factors such as initial wealth position, experience and knowledge, cognitive ability, and emotion. The higher the initial position and the emotion, the lower the sensitivity of perceiving gains and losses; the more experience, knowledge, and cognitive ability a person has, the higher his sensitivity to outcomes and probabilities. These individual factors were left uncontrolled and unexplained in past experiments. Therefore, their findings about the patterns of values and weights are at best approximate.

3. Properties of the coarse utility function

In the spectrum of risk attitudes, the maximin criterion corresponds to extreme risk aversion, while the maximax criterion corresponds to extreme risk seeking. Let U^{min} and U^{max} denote the utility functions of the maximin and maximax criteria, respectively. Let U^E denote the expected utility function. If we arrange lottery L into $\{(o_{-m}, P(o_{-m})), \ldots, (o_{-1}, P(o_{-1})), (o_0, P(o_0)), (o_1, P(o_1)), \ldots, (o_n, P(o_n))\}$ such that $o_{-m} \lesssim o_{-(m-1)} \lesssim \cdots \lesssim o_{-1} \lesssim o_0 \lesssim o_1 \lesssim o_2 \lesssim \cdots \lesssim o_n$, then its rank- and sign-dependent utility is represented as follows:

$$U^{RSD}(L) = v(o_n)W^+(P(\tilde{o}_n)) + \sum_{i=0}^{n-1} v(o_i)(W^+(P(\tilde{o}_i)) - W^+(P(\tilde{o}_{i+1}))) \\ + v(o_{-\mathrm{m}})W^-(P(\underset{\sim}{o}_{-m})) + \sum_{i=0}^{-(m-1)} v(o_i)(W^-(P(\underset{\sim}{o}_i)) - W^-(P(\underset{\sim}{o}_{i-1}))). \quad (8)$$

Property 1. For any lottery L, the following inequalities hold:

(1) $U^{min}(L) \leq U^C(L) \leq U^{RSD}(L) \leq U^{max}(L)$ if L is positive;

(2) $U^{min}(L) \leq U^{RSD}(L) \leq U^C(L) \leq U^{max}(L)$ if L is negative;

(3) $U^{min}(L) \leq U^{RSD}(L) \cong U^C(L) \leq U^{max}(L)$ if L mixed.

This property states that, compared to the coarse utility function, the rank- and sign-dependent utility function overestimates the utility of a positive lottery, underestimates the utility of a negative lottery, and approximates the utility of a mixed lottery. Therefore, we can regard the coarse utility function as an approximation of the rank- and sign-dependent utility function.

Generalized models compute the utility of a lottery by aggregating the utility of each outcome-probability pair of the lottery, regardless of the value of the outcome or the belief of its likelihood. Compared to the coarse utility function, their overestimation or underestimation comes from low-valued outcomes and relatively unlikely outcomes, according to the proof of Property 1 in section 6. We hypothesize that by including outcomes with relatively low values, or outcomes with low likelihoods, generalized models bias the utilities of lotteries. For example, consider the St. Petersburg lottery {(\$2, 1/2), (\$$2^2$, $1/2^2$), ... , (\$$2^n$, $1/2^n$, ...}. For large values of n, the outcomes \$$2^n$ are extremely unlikely, and these outcomes should not contribute to the aggregate utility. The inclusion of these terms overestimates the utility of a lottery and is not desirable.

Property 2. Suppose L_1 and L_2 are positive lotteries, and $0 \leq \alpha \leq 1$. Then,

$$U^C(\{(L_1, \alpha), (L_2, 1 - \alpha)\}) \leq \alpha U^C(L_1) + (1 - \alpha)U^C(L_2) \tag{9}$$

if $W^+(p)$ is convex or linear; suppose L_1 and L_2 are negative lotteries and $0 \leq \alpha \leq 1$. Then,

$$U^C(\{(L_1, \alpha), (L_2, 1 - \alpha)\}) \geq \alpha U^C(L_1) + (1 - \alpha)U^C(L_2) \tag{10}$$

if $W^-(p)$ is convex or linear.

Note that empirical weighting functions are convex when p is large, and concave when p is small. The critical points p_c at which $W^+(p_c) = p_c$ or $W^-(p_c) = p_c$ vary from study to study. For example, Preston and Baratta (1948) found $p_c \approx 0.2$. Hogarth and Einhorn (1990), on the other hand, found that the location of p_c depends on the absolute size and the sign of payoffs. In general, $p_c < 0.5$ for $W^+(p)$ and $p_c > 0.5$ for $W^-(p)$. Specifically, the convex region of $W^+(p)$ grows as o^* increases, and the concave region of $W^-(p)$ grows as $|o_*|$ increases. A recent study in Tversky and Kahneman (1992) found $p_c \approx 0.32$ for both $W^+(p)$ and $W^-(p)$.[3]

Property 1 states the difference between the coarse utility U^C and the rank- and sign-dependent utility U^{RSD}. Property 2 states the nonlinearity of the coarse utility function and its difference from the expected utility. As we see from (8), when weighting functions are linear, the rank- and sign-dependent utility U^{RSD} reduces to the expected utility. The expected utility function has an obvious advantage over any other generalizations in evaluating multistage choices in decision trees (Raiffa, 1968), or their variants such as those which influence diagrams (Olmsted, 1983; Schacter, 1986) and valuation networks (Shenoy, 1992, 1994). According to Sarin and Wakker (1994), violations of independence are not allowed in decision tree analysis if one accepts folding back and interchangeability of consecutive event nodes. As we reviewed before, the rank- and sign-dependent utility function allows violations of independence. Shortly we will show that the coarse utility function does not observe independence in general. Therefore, the standard rollback procedure in decision-tree analysis does not work for both the rank- and sign-dependent and the coarse utility calculi. Nevertheless, we can easily show that the rank- and sign-dependent utility and the coarse utility are reconciled if the value function $v(o)$ is a step function, as suggested by Simon (1955). Furthermore, when the weighting functions are linear, the expected utility, the rank- and sign-dependent utility, and the coarse utility are all reconciled.

The coarse utility and the rank- and sign-dependent utility are also reconciled when a lottery has only one nonzero outcome. If $L = \{(0, 1 - p), (o, p)\}$, $U^C(L) = U^{RSD}(L) = v(o)W^+(P(\tilde{o}))$, if o is a gain, and $v(o)W^-(P(\underset{\sim}{o}))$ if o is a loss. This simple property can be utilized to derive a procedure for assessing value or weighting functions. In the expected utility theory, a value function is elicited from a DM through a procedure that assesses certainty equivalents of standard lotteries. In the assessment procedure, given a standard lottery, the DM finds a certain outcome such that the standard lottery is equally preferred to the certain outcome. The same procedure will also work for our coarse utility function. For a degenerate lottery $L = \{(o, 1)\}$, the utility is $U^C(L) = v(o)$. In words, the value of an outcome is the utility of the corresponding degenerate lottery. For a positive standard lottery $L_s^+ = \{(0, 1 - p), (o^*, p)\}$, where o^* is a best outcome, $U^C(L_s^+) = v(o^*)W^+(p) = W^+(p)$. Suppose the certainty equivalent of L_s^+ is o, which is the degenerate lottery $L = \{(o, 1)\}$. Then, $U^C(L) = v(o)$. Therefore $U^C(L_s^+) = W^+(p) = U^C(L) = v(o)$. In words, the weight of the best outcome in the standard lottery is the value of its certainty equivalent. Similarly, for a negative standard lottery $L_s^- = \{(0, 1 - p), (o_*, p)\}$, where o_* is a worst loss, suppose its certainty equivalent is o. Then $W^-(p) = -v(o)$, i.e., the weight of the worst loss in the standard lottery is the negative value of its certainty equivalent. Therefore, when decision weights are known, values can be assessed using the standard lottery approach, and vice versa.

Property 3. The underlying preference relation of the coarse utility function is asymmetric, negatively transitive, and continuous.

Property 4. If L_1 stochastically dominates L_2, then $U^C(L_1) \geq U^C(L_2)$, and for any L_3 and λ,

$$U^C(\{(L_1, \lambda), (L_3, 1 - \lambda)\}) \geq U^C(\{(L_2, \lambda), (L_3, 1 - \lambda)\}).$$

If $U^C(L_1) \geq U^C(L_2)$ for any increasing function $v(o)$, then L_1 stochastically dominates L_2.

These two properties show the normative appeal of the coarse utility function. Property 3 relates the coarse utility function to the preference order. Property 4 states that the coarse utility function satisfies the requirement of consistency with the first-order stochastic dominance principle. Note that the dominance of L_1 over L_2 implies $U^C(L_1) \geq U^C(L_2)$. This inequality is not strict. Indeed, there are situations where L_1 stochastically dominates L_2, but $U^C(L_1) = U^C(L_2)$. For example, suppose L is a positive lottery, and o maximizes $v(o)W^+(P(\tilde{o}))$. Then, according to (4), the coarse approximation $L(o)$ and L have the same coarse utility value, while L weakly dominates $L(o)$. The coarse utility function shows its weakness in this respect. One can construct other similar counterintuitive examples. Our defense here is that the coarse utility function approximates the utility of a lottery. Thus, $U^C(L_1) = U^C(L_2)$ should be interpreted as $U(L_1) \approx U(L_2)$, and this is not inconsistent with our empirical findings. According to Property 4, the coarse utility function does not admit the violations of dominance. Therefore, this example simply shows that the coarse utility is sometimes incapable of identifying preferences among weakly differentiated lotteries.

In general, the coarse utility function does not satisfy the independence axiom of the expected utility theory. For example, if $L_1 = \{(o_2, 0.8), (o_3, 0.2)\}$, $L_2 = \{(o_1, 0.4), (o_3, 0.6)\}$, $L_3 = \{(o_1, 0.1), (o_3, 0.9)\}$, where $v(o_1) = 0.1, v(o_2) = 0.65, v(o_3) = 0.8$. Suppose $W^+(p)$ is an identity. Using (1), we can compute $U^C(L_1) = 0.65 > U^C(L_2) = 0.48$. Let $\lambda = 0.3$. Then

$$U^C(\{(L_1, \lambda), (L_3, 1 - \lambda)\}) = MAX\{0.1, 0.604, 0.552\} = 0.604,$$
$$U^C(\{(L_2, \lambda), (L_3, 1 - \lambda)\}) = MAX\{0.1, 0.526, 0.648\} = 0.648.$$

Therefore, $U^C(\{(L_1, \lambda), (L_3, 1 - \lambda)\}) < U^C(\{(L_2, \lambda), (L_3, 1 - \lambda)\})$. The independence of irrelevant alternatives is a controversial property of the expected utility theory, and one that is also violated in observed decision-making behavior. Therefore, we do not view the noncompliance of this property by the coarse utility function as a flaw.

4. Empirical evidence

Like any decision-weight model, the coarse utility function uses decision weights to reflect DMs' attitudes to probabilities. Therefore, it permits the analysis of phenomena associated with the distortion of subjective probability and accommodates some empirical violations of the expected utility theory. However, because attitudes to outcomes and attitudes to probabilities are not separable, one cannot assess values and weights independently. As we showed in section 3, by assessing the certainty equivalent of a standard lottery, we can assess weights given values or values given weights, but not both simultaneously. This difficulty is prevalent in any other theory. In the expected utility theory, weighting functions are linear. The probability of the best outcome of a standard lottery is the same as the value of the certainty equivalent (Raiffa, 1968; Keeney and Raiffa, 1976). In decision-weight theories, as exemplified by Hogarth and Einhorn (1990), Camerer and Ho (1991), and Tversky and Kahneman (1992), people assess decision weights by assuming a linear or power value function. The resulting findings confound the general test of a full-fledged, decision-weight theory with that of a specific value function. As Tversky and Kahneman (1992) have pointed out, the assessment of values and weights for a complex model, such as the rank- and sign-dependent utility theory, is problematic. More research with different perspectives is needed in this regard. For example, as a speculation, we might be able to trace values from the mechanism of their origin and formation.

In the coarse utility theory, a value function $v(o)$ is normalized such that $v(o_*) = -1$, $v(0) = 0$, and $v(o^*) = 1$. Therefore, in applying the coarse utility theory, an important step is to specify the outcome space Θ, its worst loss o_*, and its best gain o^*, which should be fixed in a given decision context because of the normalization of $v(o)$ and its impact on assessing decision weights.

In this section, we use three examples to compare the coarse utility function with the expected utility function and the rank- and sign-dependent utility function of Luce and Fishburn (1991) and Tversky and Kahneman (1992). Since the magnitude of payoffs in our examples is close to that in the experiments of Tversky and Kahneman (1992), a value function and a weighting function that fit their data, we believe, could approximately describe the choice behavior of the subjects in our examples. Tversky and Kahneman (1992) obtained such a pair of fitted functions as follows:

$$v(o) = o^{0.88} \text{ for } o \geq 0 \text{ and } W^+(p) = \frac{p^{0.61}}{[p^{0.61} + (1-p)^{0.61}]^{1/0.61}}, \tag{11}$$

which is quite close to the result of similar studies in Camerer and Ho (1991). Using Karmarkar's weighting function form (1978), we can also fit the following functions to their data very well:[4]

$$v(o) = o^{0.761} \text{ for } o \geq 0 \text{ and } W^+(p) = \frac{p^{0.55}}{p^{0.55} + (1-p)^{0.55}}. \tag{12}$$

As Tversky and Kahneman (1992) indicate, their results provide a reasonably good approximation to both aggregate and individual data only for probabilities in the range between 0.05 and 0.95. Since there are no data available to assess $W^+(p)$ for $p < 0.05$, we use interpolations to estimate $W^+(p)$ for extremely small p in Example 2. Several experiments were conducted in a research seminar and three undergraduate classes in the School of Business at the University of Kansas. In all the examples, the experimental results are robust to the choice of value and weighting functions (11) and (12), and strongly support the descriptive accuracy of the coarse utility function.

Example 1. Let $L_1 = \{(\$0, 0.1), (\$100, 0.1), (\$200, 0.1), (\$300, 0.1), (\$400, 0.1), (\$500, 0.1), (\$600, 0.1), (\$700, 0.1), (\$800, 0.1), (\$900, 0.1)\}$ and $L_2 = \{(\$350, 1)\}$. Which lottery do you prefer, L_1 or L_2? 80% of the subjects preferred L_2 to L_1. Also, when those who preferred L_2 to L_1 were asked what was the most they were willing to pay for L_1, 60% responded between \$100 and \$250, 30% responded \$300, and 10% responded less than \$100. Assume the value function and the weighting function in (11). Then, according to the expected utility theory, the certainty equivalent for L_1 is $CE^E(L_1) = \$435.58$; thus L_1 is preferred to L_2. According to the rank- and sign-dependent utility theory, the certainty equivalent for L_1 is $CE^{RSD}(L_1) = \$363.34$; thus L_1 is also preferred to L_2. According to the coarse utility theory, the certainty equivalent for L_1 is $CE^C(L_1) = \$192.45$; thus L_2 is preferred to L_1. Similarly, assume (12). Then, the expected utility theory suggests that the certainty equivalent for L_1 is $CE^E(L_1) = \$419.4$; thus, L_1 is preferred to L_2. According to the rank- and sign-dependent utility theory, the certainty equivalent for L_1 is $CE^{RSD}(L_1) = \$399.54$; thus L_1 is also preferred to L_2. According to the coarse utility theory, the certainty equivalent for L_1 is $CE^C(L_1) = \$209.5$; thus L_2 is preferred to L_1.

Example 2. Consider the truncated St. Petersburg lottery, $L = \{(\$2, 1/2), (\$4, 1/4), \ldots, (\$2^{10}, 1/2^{10}), (\$2^{11}, 1/2^{10})\}$. We first assume the value function (11) and normalize it by dividing each value by $2048^{0.88}$. We compute $W^+(p)$ for $p > 0.05$ using (11). We estimate $W^+(p)$ for $p < 0.05$ based on the interpolation formula: $W^+(p) = 2.36p$ for $p < 0.05$. The resulting values and weights are shown in table 1. Then, according to the expected utility theory, the certainty equivalent for L is $CE^E(L) = \$9.59$; according to the rank- and sign-dependent utility theory, the certainty equivalent for L is $CE^{RSD}(L) = \$15.99$; according to the coarse utility theory, the certainty equivalent for L is $CE^C(L) = \$3.63$. We then repeat the above calculations using (12). We find the certainty equivalent for L by the expected utility theory is $CE^E(L) = \$7.54$; by the rank- and sign-dependent utility theory is $CE^{RSD}(L) = \$17.04$; and by the coarse utility theory is $CE^C(L) = \$3.51$. In our experiments, 88% of the subjects were willing to pay no more than \$8 for the gamble, 12% were willing to pay between \$10 and \$16, 72% between \$2 and \$4, and 61% between \$3 and \$4. The results support the prediction of the coarse utility theory.

Example 3. Consider a variant of Allais's Paradox (Kahneman and Tversky, 1979). Consider a pair of lotteries $L_1 = \{(\$3000, 1)\}$, and $L_2 = \{(\$4000, 0.8), (\$0, 0.2)\}$, and another pair $L_3 = \{(\$3000, 0.25), (\$0, 0.75)\}$ and $L_4 = \{(\$4000, 0.2), (\$0, 0.8)\}$. According to an experiment conducted by Kahneman and Tversky, many subjects preferred L_1 to L_2, and L_4 to L_3. Let $v(0) = 0$. Then the expected utility theory will produce the inconsistency that $0.8 < v(3000)/v(4000) < 0.8$. Using the coarse utility function, the choices of majority in the experiment can be represented by

$$U^C(L_1) = v(3000) > U^C(L_2) = v(4000)W^+(0.8),$$
$$U^C(L_3) = v(3000)W^+(0.25) < U^C(L_4) = v(4000)W^+(0.2).$$

Then we have the inequality that $W^+(0.8) < v(3000)/v(4000) < W^+(0.2)/W^+(0.25)$. This result is consistent with the experimental finding that people tend to overweight small probabilities and underweight large ones (Kahneman and Tversky, 1979; Hogarth and Einhorn, 1990; Tversky and Kahneman, 1992). For example, one can verify that it is consistent with the empirical curves shown in (11) and (12). In fact, when a lottery has only one nonzero outcome or has one positive and one negative outcome, the coarse utility function and the prospect theory are reconciled. Any empirical evidence with such kinds of lotteries will equally support both the coarse utility theory and the prospect theory.

Table 1. The value function and the weighting function of form (11) for example 2

o	$\$2^{11}$	$\$2^{10}$	$\$2^9$	$\$2^8$	$\$2^7$	$\$2^6$	$\$2^5$	$\$2^4$	$\$2^3$	$\$2^2$	\$2
$v(o)$	1	0.543	0.295	0.160	0.087	0.047	0.026	0.014	0.008	0.004	0.002
$P(\bar{o})$	$1/2^{10}$	$1/2^9$	$1/2^8$	$1/2^7$	$1/2^6$	$1/2^5$	$1/2^4$	$1/2^3$	$1/2^2$	1/2	1
$W^+(p)$	.0023	.0036	.0093	.0185	.0370	.0740	.1475	.2077	.2907	.4206	1.0

5. Conclusion

In this article, we argue that DMs use approximations in evaluating complex lotteries. We describe a coarse approximation scheme which approximates a many-outcome lottery by several binary lotteries. We define the utility of a lottery to be the utility of its best binary approximation. We thus propose a new descriptive model called the coarse utility function. The coarse utility function is shown to be descriptive. It accommodates violation of independence evidenced by Allais's paradox and it provides a better resolution to the St. Petersburg paradox than the expected and the rank- and sign-dependent utility theories. The coarse utility theory also has many prescriptive advantages for practical decision making. It requires fewer assessments than any other exact theories. It is flexible enough to allow both additive and nonadditive probabilities and flexible judgmental information. It generalizes the mean-variance model for portfolio selection (Liu, 1994). The coarse utility function has been established using a mixed rational and plausible approach. Despite its descriptive focus, it has some normative appeal. The model is shown to underlie a continuous and weak preference relation. It satisfies the requirement of consistency with the first-order stochastic dominance principle. But the underlying preference relation is independent of irrelevant alternatives only when the preference relation is implied by the stochastic dominance.

Like any rank-dependent utility models, the coarse utility function is generally nonlinear with respect to probabilities. Like the rank- and sign-dependent utility function of Luce and Fishburn (1991, 1994) and Tversky and Kahneman (1992), the coarse utility function allows different decision weights for gains and losses. Thus, it permits the analysis of phenomena associated with the distortion of probabilities and the framing effect. However, there are some differences between the coarse utility function and the rank-dependent utility theories. First, the coarse utility function is not based on weighted averages. The utility of a lottery is the utility of its best approximation, where probabilities (or distorted probabilities) act as discount rates. In a certain sense, the coarse utility function can be regarded as an approximation of the rank- and sign-dependent utility function of Luce and Fishburn (1991) and Tversky and Kahneman (1992). When a lottery is a regular prospect (Kahneman and Tversky, 1979), or when a Simon's satisficing value function is assumed, the coarse utility theory and the rank- and sign-dependent utility theory are reconciled.

Second, in the coarse utility theory, decision weights are formalized rather than defined in an ad hoc fashion. The weighting function $W^+(p)$ for gains measures the subjective value of the probability p with which the most desirable gain is obtained. The weighting function $W^-(p)$ for losses measures the subjective value of the probability p with which the most undesirable loss is avoided. This formalization allows a natural interpretation of overweighting small probabilities and underweighting large ones, as well as the value $\times$ probability interaction. The formalization also implies that the weighting functions depend only on the decision context characterized by the maximum gain and loss. However, in rank-dependent theories, weighting functions are assumed to be both context- and lottery-dependent. As Quiggin (1993) has elaborated, the overweighting of small probabilities should be applied to low-probability extreme outcomes, and not to low-probability intermediate outcomes. Thus, the same cumulative probability may be weighted differently in different lotteries. This feature makes the empirical tests and prescriptive applications of the rank-dependent theories difficult. Also, the weighting functions formalized in this article, have no constraints such as $W(0.5) = 0.5$ or $W(p)$ is convex or concave, etc., unlike some earlier decision-weight models. Also, unlike several previous sign-dependent theories (Kahneman and Tversky, 1979; Starmer and Sugden, 1989; Schmeidler, 1989; Gilboa, 1987; Nakamura, 1990; Wakker, 1989), we do not assume $W^+(p) = W^-(p)$ or $W^+(p) = 1 - W^-(1 - p)$.

Third, in the expected utility theory, Arrow (1971) and Pratt (1964) show the equivalence between risk aversion and the concavity of a value function. This equivalence is not sustained in more general theories because of the additional dimension of attitudes toward probability preferences. As Friedman and Savage (1948) have argued, risk-seeking behavior can be modeled in terms of probability attitudes, without the use of convex segments of a value function. In rank-dependent utility theories, when a value function is linear, Yaari (1987) shows that a DM is pessimistic if and only if he/she always overweights probabilities and that a DM is optimistic if and only if he/she always underweights probabilities. If a value function is assumed to be concave, Quiggin (1993) proves that a DM is risk averse if he/she always overweights probabilities. In general, Chew, Karni, and Safra (1987) establish that a rank-dependent utility model follows second-order stochastic dominance if and only if both a value function and a weighting function are concave. In the coarse utility theory, Liu (1994) establishes a measure of risk attitudes that resembles the Arrow–Pratt measure. He shows that a DM with a concave value function and a convex weighting function is more conservative than a DM with a linear value function and a linear weighting function. A DM with a convex value function and a concave weighting function is more risk seeking than a DM with a linear value function and a linear weighting function.

We agree with Tversky and Kahneman (1992) that theories of choice are at best approximate and incomplete. Besides its incapability of identifying weakly differentiated preferences, the coarse utility function is also not general enough to accommodate violations of transitivity (Grether and Plott, 1979; Slovic and Lichtenstein, 1983) and violations of dominance (Coombs, 1975). The coarse utility is subject to this criticism in company with many other utility theories such as the rank- and sign-dependent and expected utility functions. Quiggin (1982) suggests several answers to this objection. According to Fishburn (1988), most theorists regard the reduction principle, asymmetry of strict preference, and first-degree stochastic dominance as normatively essential, and there is little concern about possible failures of the continuity axiom. The view that transitivity can no longer be regarded as a tenet of the normative creed is presently a minority position.

6. Proofs

Proof of Theorem 1. We will only prove the sufficiency part of the theorem, because the necessity is obvious. Under Axioms 1-3, Debreu (1954) shows that there is a real, continuous, and order-preserving function V on $\Theta \times [0, 1]$. V is unique up to a strictly increasing transformation. Let $x = V(o, 1)$ and $y = V(o^*, p)$. Since V is continuous and both Θ and $[0, 1]$ are connected, the ranges of x and y are real intervals. Since o^* is strictly preferred to 0, $(o^*, 1)$ stochastically dominates $(0, 1)$ and $(o^*, 0)$. According to Axiom 4, $V(o^*, 1) > V(0, 1)$ and $V(o^*, 1) > V(o^*, 0)$. Therefore, the ranges of x and y are nondegenerate. Let $x_0 = V(0, 1), x^* = V(o^*, 1), y_0 = V(o^*, 0)$, and $y^* = x^* = V(o^*, 1)$. Then, on the space $[\Theta^+ \cup \{0\}] \times [0, 1]$, the range of x is $[x_0, x^*]$, and the range of y is $[y_0, y^*]$. Since for any o in $\Theta^+, o > 0$, by Axiom 4, $V(o, 1) > V(0, 1) = x_0$. Similarly, for any $p > 0$, dominance implies $V(o^*, p) > V(o^*, 0) = y_0$. Therefore, if V is restricted to the space $\Theta^+ \times (0, 1]$, the range of x is $(x_0, x^*]$, and the range of y is $(y_0, y^*]$.

For any $p > 0, (o,p) \gtrsim (o',p)$ implies that $o \gtrsim o'$, because, otherwise, if $o' > o$, Axiom 4 will imply $(o,p) < (o',p)$. Thus, for all $p > 0$, dominance further implies $(o,p) \gtrsim (o',p)$. Similarly, $(o,p) \gtrsim (o,p')$ for some $o \in \Theta^+$ implies $p \geq p'$, and thereby implies $(o,p) \gtrsim (o,p')$ for all $o \in \Theta^+$ by Axiom 4. Therefore, in terms of multiattribute decision theory (Keeney and Raiffa, 1976), Axiom 4 ensures that $o \in \Theta^+$ and $p \in (0, 1]$ are preferentially independent in the sense that

$$(o,p) \gtrsim (o',p) \text{ for any } p \in (0, 1] \text{ implies } (o,p) \gtrsim (o',p) \text{ for all } p \in (0, 1],$$
$$(o,p) \gtrsim (o,p') \text{ for any } o \in \Theta^+ \text{ implies } (o,p) \gtrsim (o,p') \text{ for all } o \in \Theta^+.$$

According to Theorem 7.1.7 in Sawaragi, Nakayama, and Tanino (1985), the preferential independence implies that $V(o,p)$ on $\Theta^+ \times (0, 1]$ can be decomposed as $V(o,p) = B(x, y)$, where B is continuous and strictly increasing in $x \in (x_0, x^*]$ and $y \in (y_0, y^*]$. The contours of the function $B(x,y)$ constitute a family of curves $\{\mathcal{T} \mid C = V(o,p), (o,p) \in \Theta^+ \times (0, 1]\}$, where

$$\mathcal{T} = \{(x,y) \mid B(x,y) = C, x \in (x_0, x^*], y \in (y_0, y^*]\}.$$

Besides this family, we can define two other families of curves, which are respectively parallel horizontal and vertical lines. At any point (x,y) in the region $G = (x_0, x^*] \times (y_0, y^*]$, exactly one curve of each family goes through it. Two curves of different families have at most one common point. Assume (x_1, y_1) and (x_2, y_2) are on the same curve $\mathcal{T}_1$, while (x_2, y_3) and (x_3, y_1) are on the same curve $\mathcal{T}_2$ (see figure 6). Let

$$x_1 = V(o_1, 1), x_2 = V(o_2, 1), x_3 = V(o_3, 1),$$
$$y_1 = V(o^*, p_1), y_2 = V(o^*, p_2), y_3 = V(o^*, p_3).$$

Then, we have $(o_1,p_1) \sim (o_2,p_2)$ and $(o_2,p_3) \sim (o_3,p_1)$, which joint imply $(o_1,p_3) \sim (o_3,p_2)$. Since (o_1,p_3) corresponds to (x_1,y_3) and (o_3,p_2) corresponds to (x_3,y_2), $(o_1,p_3) \sim (o_3,p_2)$ implies that the points (x_1,y_3) and (x_3,y_2) are on the same indifference curve. Therefore, according to Blaschke (1928), there exists a topological transformation that transforms all the indifference curves into parallel straight lines. In other words, there exist continuous and strictly increasing functions $f(x)$ and $g(y)$, which are unique up to a positive linear transformation, such that each indifference curve is a straight line: $f(x) + g(y) = C$. We impose that $f(x^*) = 0$ and $g(y^*) = 0$. Then $f(x)$ and $g(y)$ are determined up to a

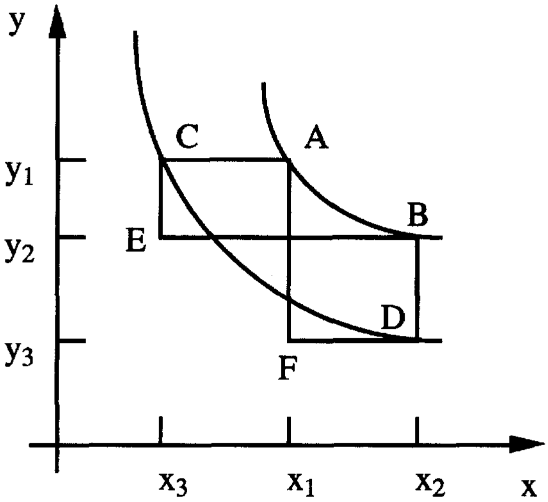

Figure 6. Verifying Thomsen-Blaschke condition.

positive affine transformation. If we let $F(x) = e^{f(x)}$ and $G(y) = e^{g(y)}$, then any indifference curve in $(x_0, x^*] \times (y_0, y^*]$ can be represented by $F(x)G(y) = C$. F and G are both strictly increasing functions of x and y. $\hat{F}(o, 1) = F(V(o, 1))$, $\hat{G}(o^*, p) = G(V(o^*, p))$, $\hat{V}(o, p) = \hat{F}(o, 1)\hat{G}(o^*, p)$. Then, obviously, $\hat{V}(o, p)$ is an order-preserving function on $\Theta^+ \times (0, 1]$. Since $F(x^*) = G(y^*) = 1$, it is easy to check that $\hat{V}(o^*, 1) = 1$, $\hat{V}(o, 1) = \hat{F}(o, 1)$, and $\hat{V}(o^*, p) = \hat{G}(o^*, p)$. Thus, $\hat{V}(o, p)$ satisfies

$$\hat{V}(o,p) = \hat{V}(o,1)\hat{V}(o^*,p) \text{ for any } (o,p) \in \Theta^+ \times (0,1]. \tag{13}$$

$\hat{V}$ is unique in the sense that, if V' is another order-preserving function satisfying (13), then $V' = \hat{V}^\alpha$ where $\alpha > 0$. Finally, for $(o, p) \in \{0\} \times [0, 1]$ or $\Theta^+ \times \{0\}$, we define

$$\hat{V}(0,p) = \hat{V}(o,0) = 0 \text{ for any } o \in \Theta^+ \text{ and } p \in [0,1].$$

Then, $\hat{V}(o, p)$ is an order-preserving function and satisfies $\hat{V}(o, p) = \hat{V}(o, 1)\hat{V}(o^*, p)$ on $[\Theta^+ \cup \{0\}] \times [0, 1]$. The uniqueness of $\hat{V}$ is unchanged. □

Proof of Theorem 2. The proof of Theorem 1 can be slightly modified as a proof of Theorem 2. Under Axioms 1-3, there is a real, continuous, and order-preserving function V on $\Theta \times [0, 1]$. V is unique up to a strictly increasing transformation. Let $x = V(o, 1)$ and $y = V(o_*, p)$. Since 0 is strictly preferred to o_*, both (0, 1) and o_*, 0) stochastically dominate $(o_*, 1)$. According to Axiom 4, $V(o_*, 1) < V(0, 1)$, and $V(o_*, 1) < V(o_*, 0)$. Let $x_0 = V(0, 1)$, $x_* = V(o_*, 1)$, $y_0 = V(o_*, 0)$, and $y_* = x_* = V(o_*, 1)$. Then, on $[\Theta^- \cup \{0\}] \times [0, 1]$, the range of x is $[x_*, x_0]$, and the range of y is $[y_*, y_0]$. Since for any o in Θ^-, $o < 0$, by Axiom 4, $V(o, 1) < V(0, 1) = x_0$. Similarly, for any $p > 0$, dominance implies $V(o_*, p) < V(o_*, 0) = y_0$. Therefore, if V is restricted to the space $\Theta^- \times (0, 1]$, the range of x is $[x_*, x_0)$, and the range of y is $[y_*, y_0)$.

Axiom 4 ensures that $o \in \Theta^-$ and $p \in (0, 1]$ are preferentially independent. Thus, $V(o, p)$ on $\Theta^- \times (0, 1]$ can be decomposed as $V(o, p) = B(x, y)$, where B is continuous and strictly increasing in $x \in [x_*, x_0)$ and $y \in [y_*, y_0)$. The contours of the function $B(x, y)$ constitute a family of curves $\{\mathcal{T} \mid C = V(o, p), o, p) \in \Theta^- \times (0, 1]\}$, where

$$\mathcal{T} = \{(x, y) / B(x, y) = C, x \in [x_*, x_0), y \in [y_*, y_0)\}.$$

According to Blaschke (1928), there exists a topological transformation that transforms all the indifference curves into parallel straight lines. In other words, there exist strictly increasing functions $f(x)$ and $g(y)$, which are unique up to a positive linear transformation, such that each indifference curve is transformed into a straight line: $f(x) + g(y) = C$, which can also be written as $-f(x) - g(y) = C$. We impose that $f(x_*) = 0$ and $g(y_*) = 0$. Then $f(x)$ and $g(y)$ are determined up to a positive affine transformation. If we let $F(x) = e^{-f(x)}$ and $G(y) = e^{-g(y)}$, then any indifference curve in $[x_*, x_0) \times [y_*, y_0)$ can be represented by $F(x)G(y) = C$. F and G are both strictly decreasing functions of x and y. Let $\hat{F}(o, 1) = -F(V(o, 1))$, $\hat{G}(o_*, p) = -G(V(o_*, p))$, $\hat{V}(o, p) = -\hat{F}(o, 1)\hat{G}(o_*, p)$. Then, we can verify that $\hat{V}(o, p)$ is an order-preserving function on $\Theta^- \times (0, 1]$. Since $F(x_*) = G(y_*) = 1$, it is easy to check that $\hat{V}(o_*, 1) = -1$, $\hat{V}(o, 1) = \hat{F}(o, 1)$, and $\hat{V}(o_*, p) = \hat{G}(o_*, p)$. Thus, $\hat{V}(o, p)$ satisfies

$$\hat{V}(o, p) = -\hat{V}(o, 1)\hat{V}(o_*, p) \text{ for any } (o, p) \in \Theta^- \times (0, 1]. \tag{14}$$

$\hat{V}$ is unique in the sense that, if V' is another order-preserving function satisfying (14), then $V' = -(-\hat{V})^\alpha$ where $\alpha > 0$. Finally, for $(o, p) \in \{0\} \times [0, 1]$ or $\Theta^- \times \{0\}$, we define

$$\hat{V}(0, p) = \hat{V}(o, 0) = 0 \text{ for any } o \in \Theta^- \text{ and } p \in [0, 1].$$

Then, $\hat{V}(o, p)$ is an order-preserving function and satisfies $\hat{V}(o, p) = -\hat{V}(o, 1)\hat{V}(o_*, p)$ on $[\Theta^- \cup \{0\}] \times [0, 1]$. The uniqueness of $\hat{V}$ is unchanged. □

Proof of Property 1. It is easy to show that $U^{min}(L) \leq U^C(L)$, $U^{RSD}(L) \leq U^{max}(L)$. We only need to show the relationship between U^{RSD} and U^C. Let $L = \{(o_{-m}, P(o_{-m})), \ldots, (o_{-1}, P(o_{-1})), (o_0, P(o_0)), (o_1, P(o_1)), \ldots, (o_n, P(o_n))\}$ with $o_{-m} \lesssim o_{-(m-1)} \lesssim \cdots \lesssim o_{-1} \lesssim o_0 \lesssim o_1 \lesssim o_2 \lesssim \cdots \lesssim o_n$. Using the Abel's formula, we have

$$\begin{aligned} U^{RSD}(L) = {} & v(o_1)W^+(P(\tilde{o}_1)) + (v(o_2) - v(o_1))W^+(P(\tilde{o}_2)) + \cdots \\ & + (v(o_n) - v(o_{n-1}))W^+(P(\tilde{o}_n)) + v(o_{-1})W^-(P(\underset{\sim}{o}_{-1})) + (v(o_{-2}) \\ & - v(o_{-1}))W^-(P(\underset{\sim}{o}_{-2})) + \cdots + (v(o_{-m}) \\ & - v(o_{-(m-1)}))W^-(P(\underset{\sim}{o}_{-m})). \end{aligned}$$

Suppose k and j are such that $U^C(L^+) = v(o_k)W^+(P(\tilde{o}_k)$ and $U^C(L^-) = v(o_{-j})W^-(P(\underset{\sim}{o}_{-j}))$. Then, we see that $U^{RSD}(L) - U^C(L)$ is the sum of the utility values from low-valued outcomes (LV) and from relatively unlikely outcomes (LP), where

$$LV = \sum_{i=1}^{k-1} v(o_i)(W^+(P(\tilde{o}_i)) - W^+(P(\tilde{o}_{i+1}))) + \sum_{i=-1}^{-(j-1)} v(o_i)(W^-(P(\underset{\sim}{o}_i)) - W^-(P(\underset{\sim}{o}_{i-1})))$$

$$LP = (v(o_{k+1}) - v(o_k))W^+(P(\tilde{o}_{k+1})) + \cdots + (v(o_n) - v(o_{n-1}))W^+(P(\tilde{o}_n)) + (v(o_{-(j+1)}) - v(o_{-j}))W^-(P(\underset{\sim}{o}_{-(j+1)})) + \cdots + (v(o_{-m}) - v(o_{-(m-1)}))W^-(P(\underset{\sim}{o}_{-m})).$$

It is easy to observe that $U^{RSD}(L) - U^C(L) \geq 0$ if L has no negative outcomes, and $U^{RSD}(L) - U^C(L) \leq 0$ if L has no positive outcomes. However, if L is a mixed lottery, we are uncertain about the sign of $U^{RSD}(L) - U^C(L)$. Since the undervalued part for the positive decomposition and the overvalued part for the negative decomposition cancel out partially or fully, $U^C(L)$ becomes closer to $U^{RSD}(L)$ than for positive and negative lotteries. □

Proof of Property 2. Let $L_1 = \{(o_1, p_1), (o_2, p_2), \ldots, (o_n, p_n)\}$, and $L_2 = \{(o_1, q_1), (o_2, q_2), \ldots, (o_n, q_n)\}$. Assume $o_1 \lesssim o_2 \lesssim \cdots \lesssim o_n$. Then $P(\tilde{o}_k) = \sum_{i \geq k} p_i$, and $Q(\tilde{o}_k) = \sum_{i \geq k} q_i$.

$$\begin{aligned}
&\{(L_1, \alpha), (L_2, 1 - \alpha)\} \\
&\quad = \{(o_1, \alpha p_1 + (1 - \alpha)q_1), (o_2, \alpha p_2 + (1 - \alpha)q_2), \ldots, (o_1, \alpha p_n + (1 - \alpha)q_n)\}. \\
&U^C(\{(L_1, \alpha), (L_2, 1 - \alpha)\}) \\
&\quad = MAX\{W^+[\alpha P(\tilde{o}_k) + (1 - \alpha)Q(\tilde{o}_k)]v(o_k) \mid k = 1, 2, \ldots, n\} \\
&\quad \leq MAX\{\alpha W^+(P(\tilde{o}_k))v(o_k) + (1 - \alpha)W^+(Q(\tilde{o}_k))v(o_k) \mid k = 1, 2, \ldots, n\} \\
&\quad \leq \alpha\, MAX\{W^+(P(\tilde{o}_k))v(o_k) \mid k = 1, 2, \ldots, n \\
&\qquad + (1 - \alpha)MAX\{W^+(Q(\tilde{o}_k))v(o_k) \mid k = 1, 2, \ldots, n\} \\
&\quad = \alpha U^C(L_1) + (1 - \alpha)U^C(L_2).
\end{aligned}$$

Following the same strategy, we can prove (10) for negative lotteries. □

Proof of Property 3. The asymmetry and negative transitivity are obvious. In the following, we only prove the continuity. Assume $o_{-m} \lesssim o_{-(m-1)} \lesssim \cdots \lesssim o_{-1} \lesssim o_0 \lesssim o_1 \lesssim o_2 \lesssim \cdots \lesssim o_n$. Let

$$\begin{aligned}
L_1 &= \{(o_{-m}, p_{-m}), \ldots, (o_{-1}, p_{-1}), (o_0, p_0), (o_1, p_1), \ldots, (o_n, p_n)\}, \\
L_2 &= \{(o_{-m}, q_{-m}), \ldots, (o_{-1}, q_{-1}), (o_0, q_0), (o_1, q_1), \ldots, (o_n, q_n)\}, \\
L_3 &= \{(o_{-m}, r_{-m}), \ldots, (o_{-1}, r_{-1}), (o_0, r_0), (o_1, r_1), \ldots, (o_n, r_n)\},
\end{aligned}$$

$P(\tilde{o}_k) = \sum_{i \geq k} p_i$, $Q(\tilde{o}_k) = \sum_{i \geq k} q_i$, $R(\tilde{o}_k) = \sum_{i \geq k} r_i$, $P(\underset{\sim}{o}_{-k}) = \sum_{i \leq -k} p_i$, $Q(\underset{\sim}{o}_{-k}) = \sum_{i \leq -k} q_i$, and $R(\underset{\sim}{o}_{-k}) = \sum_{i \leq -k} r_i$. Suppose $U^C(L_1) = W^+(P(\tilde{o}_{k_1}))v(o_{k_1}) + W^-(P(\underset{\sim}{o}_{-j_1}))v(o_{-j_1})$, $U^C(L_3) = W^+(Q(\tilde{o}_{k_3}))v(o_{k_3}) + W^-(Q(\underset{\sim}{o}_{-j_3}))v(o_{-j_3})$, $L_1 \succ L_2$ and $L_2 \succ L_3$. We first prove that there is α such that $U^C(\{(L_1, \alpha), (L_3, 1 - \alpha)\}) > U^C(L_2)$. Set $0 < 2\varepsilon < U^C(L_1) - U^C(L_2)$. According to Theorems 1 and 2, $W^+(p)$ and $W^-(p)$ are continuous on the compact interval [0, 1]. Thus, $W^+(p)$ and $W^-(p)$ are uniformly continuous. Therefore, there is α such that, for all k and j,

$$-\varepsilon < W^+[\alpha P(\tilde{o}_k) + (1-\alpha)R(\tilde{o}_k)] - W^+(P(\tilde{o}_k)) < \varepsilon,$$
$$-\varepsilon < W^-[\alpha P(\underset{\sim}{o}_{-j}) + (1-\alpha)R(\underset{\sim}{o}_{-j})] - W^-(P(\underset{\sim}{o}_{-j})) < \varepsilon.$$

Thus

$$\begin{aligned}
&MAX\{W^+[\alpha P(\tilde{o}_k) + (1-\alpha)R(\tilde{o}_k)]v(o_k) \mid k = 1, 2, \ldots, n\} \\
&\quad \geq MAX\{W^+[P(\tilde{o}_k)]v(o_k) \mid k = 1, 2, \ldots, n\} - \varepsilon = W^+(P(\tilde{o}_{k_1}))v(o_{k_1}) - \varepsilon, \\
&MIN\{W^-[\alpha P(\underset{\sim}{o}_{-j}) + (1-\alpha)R(\underset{\sim}{o}_{-j})]v(o_{-j}) \mid j = 1, 2, \ldots, m\} \\
&\quad \geq MIN\{W^-[P(\underset{\sim}{o}_{-j})]v(o_{-j}) \mid j = 1, 2, \ldots, m\} - \varepsilon \\
&\quad = W^-(P(\underset{\sim}{o}_{-j_1}))v(o_{-j_1}) - \varepsilon, \\
&U^C(\{(L_1, \alpha), (L_3, 1-\alpha)\}) \geq U^C(L_1) - 2\varepsilon > U^C(L_2).
\end{aligned}$$

Then we prove that there is β such that $U^C(\{(L_1, \beta), (L_3, 1-\beta)\}) < U^C(L_2)$. Set

$$0 < 2\varepsilon < U^C(L_2) - U^C(L_3).$$

The uniform continuity of $W^+(p)$ and $W^-(p)$ implies that there is β such that, for all k and j,

$$-\varepsilon < W^+[\beta P(\tilde{o}_k) + (1-\beta)R(\tilde{o}_k)] - W^+(R(\tilde{o}_k)) < \varepsilon,$$
$$-\varepsilon < W^-[\beta P(\underset{\sim}{o}_{-j}) + (1-\beta)R(\underset{\sim}{o}_{-j})] - W^-(R(\underset{\sim}{o}_{-j})) < \varepsilon.$$

Therefore,

$$\begin{aligned}
&MAX\{W^+[\beta P(\tilde{o}_k) + (1-\beta)R(\tilde{o}_k)]v(o_k) \mid k = 1, 2, \ldots, n\} \\
&\quad \leq MAX\{W^+[R(\tilde{o}_k)]v(o_k) \mid k = 1, 2, \ldots, n\} + \varepsilon = W^+(R(\tilde{o}_{k_3}))v(o_{k_3}) + \varepsilon, \\
&MIN\{W^-[\beta P(\underset{\sim}{o}_{-j}) + (1-\beta)R(\underset{\sim}{o}_{-j})]v(o_{-j}) \mid j = 1, 2, \ldots, m\} \\
&\quad \leq MIN\{W^-[R(\underset{\sim}{o}_{-j})]v(o_{-j}) \mid j = 1, 2, \ldots, m\} + \varepsilon \\
&\quad = W^-(R(\underset{\sim}{o}_{-j_3}))v(o_{-j_3}) + \varepsilon, \\
&U_C(\{(L_1, \beta), (L_3, 1-\beta)\}) \leq U^C(L_3) + 2\varepsilon < U^C(L_2). \qquad \square
\end{aligned}$$

Proof of Property 4. Assume $o_{-m} \lesssim o_{-(m-1)} \lesssim \cdots \lesssim o_{-1} \lesssim o_0 \lesssim o_1 \lesssim o_2 \lesssim \cdots \lesssim o_n$. Let

$$\begin{aligned}
L_1 &= \{(o_{-m}, p_{-m}), \ldots, (o_{-1}, p_{-1}), (o_0, p_0), (o_1, p_1, \ldots, o_n, p_n)\}, \\
L_2 &= \{(o_{-m}, q_{-m}), \ldots, (o_{-1}, q_{-1}), (o_0, q_0), (o_1, q_1, \ldots, o_n, q_n)\}, \\
P(\tilde{o}_k) &= \sum_{i \geq k} p_i,\ Q(\tilde{o}_k) = \sum_{i \geq k} q_i,\ P(\underset{\sim}{o}_{-k}) = \sum_{i \leq -k} p_i, \text{ and } Q(\underset{\sim}{o}_{-k}) = \sum_{i \leq -k} q_i.
\end{aligned}$$

If L_1 stochastically dominates L_2, then $P(\tilde{o}_k) \geq Q(\tilde{o}_k), P(\underset{\sim}{o}_{-j}) \leq Q(\underset{\sim}{o} - j)$ for all k and j, and $L_1 \neq L_2$. Therefore,

$$
\begin{aligned}
&MAX\{W^+(P(\tilde{o}_k))v(o_k) \mid k = 1, 2, \ldots, n\} \\
&\qquad \geq MAX\{W^+(Q(\tilde{o}_k))v(o_k) \mid k = 1, 2, \ldots, n\}, \\
&MIN\{W^-(P(\underset{\sim}{o}_{-j}))v(o_{-j}) \mid j = 1, 2, \ldots, m\} \\
&\qquad \geq MIN\{W^-(Q(\underset{\sim}{o}_{-j}))v(o_{-j}) \mid j = 1, 2, \ldots, m\}.
\end{aligned}
$$

Therefore, $U^C(L_1) \geq U^C(L_2)$.

Let $L_3 = \{(o_{-m}, r_{-m}), \ldots, (o_{-1}, r_{-1}), (o_0, r_0), (o_1, r_1), \ldots, (o_n, r_n)\}$, $R(\tilde{o}_k) = \sum_{i \geq k} r_i$, and $R(\underset{\sim}{o}_{-k}) = \sum_{i \leq -k} r_i$. For any $\lambda \in (0, 1)$, and any k and j,

$$
\begin{aligned}
&\lambda P(\tilde{o}_k) + (1 - \lambda)R(\tilde{o}_k) \geq \lambda Q(\tilde{o}_k) + (1 - \lambda)R(\tilde{o}_k), \\
&\lambda P(\underset{\sim}{o}_{-j}) + (1 - \lambda)R(\underset{\sim}{o}_{-j}) \leq \lambda Q(\underset{\sim}{o}_{-j}) + (1 - \lambda)R(\underset{\sim}{o}_{-j}).
\end{aligned}
$$

Thus

$$
\begin{aligned}
&MAX\{W^+[\lambda P(\tilde{o}_k) + (1 - \lambda)R(\tilde{o}_k)]v(o_k) \mid k = 1, 2, \ldots, n\} \\
&\qquad \geq MAX\{W^+[\lambda Q(\tilde{o}_k) + (1 - \lambda)R(\tilde{o}_k)]v(o_k) \mid k = 1, 2, \ldots, n\}, \\
&MIN\{W^-[\lambda P(\underset{\sim}{o}_{-j}) + (1 - \lambda)R(\underset{\sim}{o}_{-j})]v(o_{-j}) \mid j = 1, 2, \ldots, m\} \\
&\qquad \geq MIN\{W^-[\lambda Q(\underset{\sim}{o}_{-j}) + (1 - \lambda)R(\underset{\sim}{o}_{-j})]v(o_{-j}) \mid j = 1, 2, \ldots, m\}.
\end{aligned}
$$

Therefore, $U^C(\{(L_1, \lambda), (L_3, 1 - \lambda)\}) \geq U^C(\{(L_2, \lambda), (L_3, 1 - \lambda)\})$. Independence is proved.

Now we prove the converse. We first prove that $U^C(L_1^+) \geq U^C(L_2^+)$ and $U^C(L_1^-) \geq U^C(L_2^-)$ for any increasing function $v(o)$. Otherwise, assume there is a value function $v(o)$ such that

$$
\begin{aligned}
&MAX\{W^+[P(\tilde{o}_k)]v(o_k) \mid k = 1, 2, \ldots, n\} \\
&\qquad < MAX\{W^+[Q(\tilde{o}_k)]v(o_k) \mid k = 1, 2, \ldots, n\}.
\end{aligned}
$$

Since $U^C(L_1) \geq U^C(L_2)$, we must have

$$
\begin{aligned}
&MIN\{W^-[P(\underset{\sim}{o}_{-j})]v(o_{-j}) \mid j = 1, 2, \ldots, m\} \\
&\qquad > MIN\{W^-[Q(\underset{\sim}{o}_{-j})]v(o_{-j}) \mid j = 1, 2, \ldots, m\}.
\end{aligned}
$$

Let $v'(o_{-j}) = \alpha v(o_{-j})$ for all $j = 1, 2, \ldots, m$ and $v'(o_k) = v(o_k)$ for all $k = 1, 2, \ldots, n$, where

$$
\alpha < \frac{MAX\{W^+[Q(\tilde{o}_k)]v(o_k)\} - MAX\{W^+[P(\tilde{o}_k)]v(o_k)\}}{MIN\{W^-[P(\underset{\sim}{o}_{-j})]v(o_{-j})\} - MIN\{W^-[Q(\underset{\sim}{o}_{-j})]v(o_{-j})\}}. \tag{15}
$$

Then,

$$
\begin{aligned}
&MAX\{W^+[P(\tilde{o}_k)]v'(o_k) \mid k = 1, 2, \ldots, n\} \\
&\qquad = MAX\{W^+[P(\tilde{o}_k)]v(o_k) \mid k = 1, 2, \ldots, n\},
\end{aligned}
$$

$$
\begin{aligned}
&MAX\{W^+[Q(\tilde{o}_k)]v'(o_k) \mid k = 1, 2, \dots, n\} \\
&\quad = MAX\{W^+[Q(\tilde{o}_k)]v(o_k) \mid k = 1, 2, \dots, n\}, \\
&MIN\{W^-[P(\underset{\sim}{o}_{-j})]v'(o_{-j}) \mid j = 1, 2, \dots, m\} \\
&\quad = \alpha MIN\{W^-[P(\underset{\sim}{o}_{-j})]v(o_{-j}) \mid j = 1, 2, \dots, m\}, \\
&MIN\{W^-[Q(\underset{\sim}{o}_{-j})]v'(o_{-j}) \mid j = 1, 2, \dots, m\} \\
&\quad = \alpha MIN\{W^-[Q(\underset{\sim}{o}_{-j})]v(o_{-j}) \mid j = 1, 2, \dots, m\}.
\end{aligned}
$$

Therefore, according to (15),

$$
\begin{aligned}
&MAX\{W^+[P(\tilde{o}_k)]v'(o_k) \mid k = 1, 2, \dots, n\} \\
&\qquad + MIN\{W^-[P(\underset{\sim}{o}_{-j})]v'(o_{-j}) \mid j = 1, 2, \dots, m\}, \\
&\quad < MAX\{W^+[Q(\tilde{o}_k)]v'(o_k) \mid k = 1, 2, \dots, n\} \\
&\qquad + MIN\{W^-[Q(\underset{\sim}{o}_{-j})]v'(o_{-j}) \mid j = 1, 2, \dots, m\}\}.
\end{aligned}
$$

That is, $U^C(L_1) < U^C(L_2)$ for the function $v'(o)$. This contradicts the given condition. Similarly, we can prove that $U^C(L_1^-) \geq U^C(L_2^-)$ for any increasing function $v(o)$.

Then we prove that, if $U^C(L_1) \geq U^C(L_2)$ for any increasing function $v(o)$ and L_1 and L_2 are either both positive or both negative, L_1 stochastically dominates L_2. If L_1 and L_2 are both positive, we need to show that $P(\tilde{o}_k) \geq Q(\tilde{o}_k)$. First we prove that $P(\tilde{o}_n) \geq Q(\tilde{o}_n)$. Otherwise, assume $0 \leq P(\tilde{o}_n) < Q(\tilde{o}_n)$. Let $v(o_k) = (k - 1)\varepsilon$ when $k \leq n - 1$ and $v(o_n)$ be any positive number. Let $\varepsilon < W^+(Q(\tilde{o}_n))v(o_n)/(n - 1)$. Then

$$
\begin{aligned}
&W^+[P(\tilde{o}_k)](k - 1)\varepsilon \leq W^+[Q(\tilde{o}_n)]v(o_n), W^+[Q(\tilde{o}_k)](k - 1)\varepsilon \leq W^+[Q(\tilde{o}_n)]v(o_n), \\
&U^C(L_1) = MAX\{W^+[P(\tilde{o}_k)](k - 1)\varepsilon, W^+[P(\tilde{o}_n)]v(o_n) \mid k = 1, 2, \dots, n - 1\} \\
&\qquad < W^+[Q(\tilde{o}_n)]v(o_n) = MAX\{W^+[Q(\tilde{o}_k)](k - 1)\varepsilon, \\
&\qquad W^+[Q(\tilde{o}_n)]v(o_n) \mid k = 1, 2, \dots, n - 1\} = U^C(L_2).
\end{aligned}
$$

It contradicts the condition that $U^C(L_2) \leq U^C(L_1)$ for any $v(o)$. Suppose for all $k \geq i + 1$, $P(\tilde{o}_k) \geq Q(\tilde{o}_k)$. We then prove that $P(\tilde{o}_i) \geq Q(\tilde{o}_i)$. Otherwise, assume $P(\tilde{o}_i) < Q(\tilde{o}_i)$. Then, if $k > i$,

$$
\begin{aligned}
&Q(\tilde{o}_k) \leq P(\tilde{o}_k) \leq P(\tilde{o}_i) < Q(\tilde{o}_i), \\
&(n - i)W^+[Q(\tilde{o}_i)] - (n - k)W^+[P(\tilde{o}_k)] > 0, \\
&(n - i)W^+[Q(\tilde{o}_i)] - (n - k)W^+[Q(\tilde{o}_k)] > 0.
\end{aligned}
$$

Let $v(o_k) = (k - 1)\varepsilon$ when $k \leq i - 1$, $v(o_k) = v(o_n) - (n - k)\varepsilon$ when $k \geq i$, and $v(o_n)$ be any positive number. Set

$$
\begin{aligned}
0 < \varepsilon < MIN\Bigg\{ &\frac{W^+[Q(\tilde{o}_i)]v(o_n)}{i - 1 + (n - i)W^+[Q(\tilde{o}_i)]}, \frac{\{W^+[Q(\tilde{o}_i)] - W^+[P(\tilde{o}_k)]\}v(o_n)}{(n - i)W^+[Q(\tilde{o}_i)] - (n - k)W^+[P(\tilde{o}_k)]}, \\
&\frac{\{W^+[Q(\tilde{o}_i)] - W^+[Q(\tilde{o}_k)]\}v(o_n)}{(n - i)W^+[Q(\tilde{o}_i)] - (n - k)W^+[Q(\tilde{o}_k)]} \;\Big|\; k > i \Bigg\}.
\end{aligned}
$$

Then

$$
\begin{aligned}
U^C(L_1) &= MAX\{W^+[P(\tilde{o}_k)](k-1)\varepsilon \text{ or } k < i, W^+[P(\tilde{o}_k)](v(o_n) \\
&\quad - (n-k)\varepsilon) \text{ for } k \geq i\} \\
&< W^+[Q(\tilde{o}_i)](v(o_n) - (n-i)\varepsilon) \\
&= MAX\{W^+[Q(\tilde{o}_k)](k-1)\varepsilon \text{ for } k < i, W^+[Q(\tilde{o}_k)](v(o_n) \\
&\quad - (n-k)\varepsilon) \text{ for } k \geq i\} = U^C(L_2).
\end{aligned}
$$

This contradicts the condition that $U^C(L_2) \leq U^C(L_1)$ for any $v(o)$. By the induction principle, we know that $P(\tilde{o}_k) \geq Q(\tilde{o}_k)$ for all $k = 1, 2, \ldots, n$. Following the same strategy as above, we can prove $P(\underset{\sim}{o}_{-j}) \leq Q(\underset{\sim}{o}_{-j})$ for any j if $U^C(L_2) \leq U^C(L_1)$ for any $v(o)$ and L_1 and L_2 are negative.

If L_1 and L_2 are mixed lotteries, as proved before, $U^C(L_1) \geq U^C(L_2)$ for any $v(o)$ implies $U^C(L_1^+) \geq U^C(L_2^+)$ and $U^C(L_1^-) \geq U^C(L_2^-)$ for any increasing function $v(o)$, which implies, in turn, $P(\tilde{o}_k) \geq Q(\tilde{o}_k)$ for all k and $P(\underset{\sim}{o}_{-j}) \leq Q(\underset{\sim}{o}_{-j})$ for any j. Thus, L_1 dominates L_2. □

Acknowledgments

This research is supported in part by the Harper Fund and in part by the National Science Foundation under Grant No. SES-9213558. The authors would like to thank Ward Edwards, Peter Fishburn, Paul Koch, Diane Lander, Hong Xu, Po-Lung Yu, and three anonymous reviewers for their comments and discussions.

Notes

1. We thank an anonymous referee who brought the interesting work of Lopes (1984, 1987) to our attention.
2. Some theorists including Tversky and Kahneman believe that normative theories will never be descriptively accurate because of empirical violations of first-order stochastic dominance and the reduction principle.
3. Figure 3 in Tversky and Kahneman (1992), which shows $p_c \approx 0.4$, does not match their stated regression results. As an anonymous referee pointed out, they have mislabeled the vertical axis of the figure.
4. The regression result in (12) was suggested by an anonymous referee.

References

Arrow, K. J. (1971). *Essays in the Theory of Risk-Bearing*. Amsterdam: North-Holland.

Bell, D. E. (1982). "Regret in Decision Making under Uncertainty." *Operations Research* 30, 961–81.

Bell, D. E., H. Raiffa, and A. Tversky. (1988). "Descriptive, Normative, and Prescriptive Interactions in Decision Making." In D. E. Bell, H. Raiffa, and A. Tversky (eds.), *Decision Making—Descriptive, Normative, and Prescriptive Interactions*. New York: Cambridge University Press, pp. 9–30.

Blaschke, W. (1928). "Topologische Fragen der Differentialgeometrie" *Mathematische Zeitschrift* 28, 150–157.

Brockett, P. L., and Y. Kahane. (1992). "Risk, Return, Skewness and Preference," *Management Science* 38, 851–866.

Brunswik, E. (1952). "The Conceptual Framework of Psychology," In *International Encyclopedia of Unified Science* 1(10). Chicago: University of Chicago Press.
Camerer, C. (1992). "Recent Tests of Generalizations of Expected Utility Theories." In W. Edwards (ed.), *Utility: Theories: Measurement and Applications*. Boston: Kluwer Academic Publishers.
Camerer, C. and T.-H. Ho. (1991). "Nonlinear Weighting of Probabilities and Violations of the Betweenness Axiom." Wharton School, University of Pennsylvania.
Chew, S. H. (1983). "A Generation of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox," *Econometrica* 51, 1065–1092.
Chew, S. H., E. Karni, and Z. Safra. (1987). "Risk Aversion in the Theory of Expected Utility with Rank-Dependent Preferences," *Journal of Economic Theory* 42, 370–381.
Cho, Y., R. D. Luce, and D. von Winterfeldt. (1993). "Tests of Assumptions about the Joint Receipt of Gambles in Rank- and Sign-Dependent Utility Theory." Research Report MBS 93-17, Institute for Mathematical Behavioral Sciences, University of California at Irvine, CA.
Cohen, M., J. Y. Jaffray, and T. Said. (1985). "Individual Behavior under Risk and under Uncertainty: An Experimental Study," *Theory and Decision* 18, 203–228.
Coombs, C. H. (1975). "Portfolio Theory and the Measurement of Risk." In M. Kaplan and S. Schwartz (eds.), *Human Judgment and Decision Processes*. New York: Academic Press, pp. 63–85.
Dawes, R. M. and B. Corrigan. (1974). "Linear Models in Decision Making," *Psychological Bulletin* 81, 97–106.
Dawes, R. M., D. Faust, and P. Meehl. (1989). "Clinical versus Actuarial Judgment," *Science* 243, 1668–1673.
Debreu, G. (1954). "Representation of a Preference Ordering by a Numerical Function." In R. M. Thrall, C. H. Coombs, and R. L. Davis (eds.), *Decision Processes*. New York: John Wiley & Sons, pp. 159–165.
Debreu, G. (1960). "Topological Methods in Cardinal Utility Theory." In K. Arrow, S. Karlin, and P. Suppes. (eds.), *Mathematical Methods in the Social Sciences*. Stanford, CA: Stanford University Press, pp. 16–26.
Edwards, W. (1955). "The Prediction of Decisions among Bets," *Journal of Experimental Psychology* 50, 201–214.
Edwards, W. (1962). "Subjective Probabilities Inferred from Decisions," *Psychological Review* 69, 109–135.
Fellner, W. (1961). "Distortion of Subjective Probabilities as Reaction to Uncertainty," *Quarterly Journal of Economics* 75, 670–689.
Fishburn, P. C. (1986). "Nontransitive Measurable Utility for Decision under Uncertainty," AT&T Bell Laboratories, Murray Hill, NJ.
Fishburn, P. C. (1988). *Nonlinear Preference and Utility Theory*. Baltimore: John Hopkins University Press.
Friedman, M., and L. J. Savage. (1948). "The Utility Analysis of Choices Involving Risk," *Journal of Political Economy* 56, 279–304.
Friedman, D. (1989). "The S-Shaped Value Function as a Constrained Optimum," *The American Economic Review* 79, 1243–1248.
Gilboa, I. (1987). "Expected Utility with Purely Subjective Non-additive Probabilities," *Journal of Mathematical Economics* 16, 65–88.
Goldstein, W. M., and H. Einhorn. (1987). "A Theory of Preference Reversals," *Psychological Review* 94, 236–254.
Grether, D. M., and C. Plott. (1979). "Economic Theory of Choice and the Preference Reversal Phenomenon," *American Economic Review* 69, 623–638.
Hammond, K. R., C. H. McClelland, and J. Mumpower. (1980). *Human Judgment and Decision Making: Theories, Methods and Procedures*. New York: Praeger.
Handa, J. (1977). "Risk, Probabilities and a New Theory of Cardinal Utility," *Journal of Political Economy* 85, 97–122.
Hershey, J. C., and P. J. H. Schoemaker. (1985). "Probability versus Certainty Equivalence Methods in Utility Measurement: Are They Equivalent?" *Management Science* 31, 1213–1231.
Hogarth, R. M., and H. J. Einhorn. (1990). "Venture Theory: A Model of Decision Weights," *Management Science* 36, 780–803.
Howard, R. (1992). "Heathens, Heretics, and Cults: The Religious Spectrum of Decision Aiding," *Interfaces* 22, 1–27.
Johnson, E. J., and J. Payne. (1985). "Effort and Accuracy in Choice," *Management Science* 31, 395–413.
Johnson, E. J., and D. Schkade. (1989). "Bias in Utility Assessments: Further Evidence and Explanations," *Management Science* 35, 406–424.

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

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

Karlin, S., and W. J. Studden. (1966). *Tchebycheff Systems: With Applications in Analysis and Statistics*. New York: Interscience.

Keefer, D. L. (1994). "Certainty Equivalents for Three-Point Discrete-Distribution Approximations," *Management Science* 40, 760-773.

Keeney, R. L., and H. Raiffa. (1976). *Decisions With Multiple Objectives: Preferences and Value Trade-offs*. New York: Wiley.

Krantz, D. H., R. D. Luce, P. Suppes, and A. Tversky. (1971). *Foundations of Measurement*. Vol. 1. New York: Academic Press.

Kunreuther, H., R. Ginsberg, L. Miller et al. (1978). *Disaster Insurance Protection: Public Policy Lessons*. New York: John Wiley & Sons.

Leland, J. W. (1986). "Individual Choice Uncertainty," Unpublished doctoral dissertation, UCLA Economics Department.

Leland, J. W. (1988). "A Theory of 'Approximate' Expected Utility Maximization." Working paper, Carnegie Mellon University Department of Social and Decision Sciences.

Liu, L. (1993). "Foundations of the Mean-Variance Analysis." Working paper 259, School of Business, University of Kansas, Kansas.

Liu, L. (1994). "Portfolio Selection Using Coarse Utilities." Working paper 260, School of Business, University of Kansas, Kansas.

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

Lopes, L. L. (1984). "Risk and Distributional Inequality," *Journal of Experimental Psychology: Human Perception and Performance* 10, 465-485.

Lopes, L. L. (1987). "Between Hope and Fear: The Psychology of Risk." In L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, Vol. 20. San Diego: Academic Press, pp. 255-295.

Luce, R. D., and P. C. Fishburn. (1991). "Rank- and Sign-Dependent Linear Utility Models for Finite First-Order Gambles," *Journal of Risk and Uncertainty* 4, 29-59.

Luce, R. D., and P. C. Fishburn. (1994). "A Note on Deriving Rank-Dependent Utility Using Additive Joint Receipts." Institute for Mathematical Behavioral Sciences, University of California at Irvine, California.

Luce, R. D., and J. W. Tukey. (1964). "Simultaneous Conjoint Measurement: A New Type of Fundamental Measurement," *Journal of Mathematical Psychology* 1, 1-27.

March, J. G. (1978). "Bounded Rationality, Ambiguity, and the Engineering of Choice," *The Bell Journal of Economics* 9, 586-608.

Markowitz, H. M. (1959). *Portfolio Selection: Efficient Diversification of Investments*. New York: John Wiley & Sons.

Nakamura, Y. (1990). "Subjective Expected Utility with Non-additive Probabilities on Finite State Space," *Journal of Economic Theory* 51, 346-366.

Olmsted, S. M. (1983). "On Representing and Solving Decision Problems." Doctoral dissertation, Department of Engineering-Economic Systems, Stanford University.

Payne, J. W. (1982). "Contingent Decision Behavior," *Psychological Bulletin* 92, 382-402.

Pratt, J. W. (1964). "Risk Aversion in the Small and in The Large," *Econometrica* 32, 122-236.

Preston, M. G., and P. Baratta. (1948). "An Experimental Study of the Auction Value of an Uncertain Outcome," *American Journal of Psychology* 61, 183-193.

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

Quiggin, J. (1993). *Generalized Expected Utility Theory: The Rank-Dependent Model*. Boston: Kluwer Academic Publishers.

Raiffa, H. (1968). *Decision Analysis*. Reading, MA: Addison-Wesley.

Roskies, R. (1965). "A Measurement Axiomatization for an Essentially Multiplicative Representation of Two Factors," *Journal of Mathematical Psychology* 2, 266-276.

Sarin, R., and P. Wakker. (1994). "Folding Back in Decision Tree Analysis," *Management Science* 40, 625-628.

Savage, L. J. (1954). *The Foundations of Statistics*. New York: John Wiley & Sons.

Sawaragi, Y., H. Nakayama, and T. Tanino. (1985). *Theory of Multiobjective Optimization*. Orlando: Academic Press.

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

Schneider, S. L., and L. L. Lopes. (1986). "Reflection in Preferences in Risk: Who and When May Suggest Why," *Journal of Experimental Psychology: Human Perception and Performance* 12, 535-548.

Schoemaker, P. J. H. (1993). "Determinants of Risk-Taking: Behavioral and Economic Views," *Journal of Risk and Uncertainty* 6, 49-73.

Schacter, R. D. (1986). "Evaluating Influence Diagrams," *Operations Research* 34, 871-882.

Shannon, C. E. (1948). "A Mathematical Theory of Communication," *Bell Systems Tech. Journal* 27, 279-423, 623-656.

Shenoy, P. P. (1992). "Valuation-Based Systems for Bayesian Decision Analysis," *Operations Research* 40. 463-484.

Shenoy, P. P. (1994). "A Comparison of Graphical Techniques for Decision Analysis." *European Journal of Operational Research* 78, 1-21.

Simon, H. A. (1955). "A Behavioral Model of Rational Choice," *Quarterly Journal of Economics* 69, 99-118.

Simon, H. A. (1978). "Rationality as Process and as Product of Thought," *Journal of American Economic Association* 68, 1-16.

Sinn, H.-W. (1967). *Economic Decision Under Uncertainty*. Amsterdam: North-Holland.

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, 645-654.

Slovic, P., and S. Lichtenstein. (1983). "Preference Reversal: A Broad Perspective," *American Economic Review* 73, 596-605.

Starmer, C., and R. Sugden. (1989). "Violations of the Independence Axiom in Common Ratio Problems: An Experimental Test of Some Competing Hypotheses," *Annals of Operations Research* 19, 79-102.

Thompson, J. (1967). *Organizations in Action*. New York: McGraw-Hill.

Tobin, J. (1959). "Liquidity Preference as Behavior towards Risk," *Review of Economic Studies* 25, 65-85.

Tversky, A., and D. Kahneman. (1992). "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty* 5, 297-323.

Tversky, A., S. Sattath, and P. Slovic. (1988). "Contingent Weighting in Judgment and Choice," *Psychological Review* 95, 371-384.

Wakker, P. P. (1989). "Continuous Subjective Expected Utility with Non-additive Probabilities," *Journal of Mathematical Economics* 18, 1-27.

Wakker, P. P., and A. Tversky. (1991). "An Axiomatization of Cumulative Prospect Theory." Unpublished manuscript, University of Nijmegan, the Netherlands.

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