Journal of Risk and Uncertainty, 8:167–196 (1994)

Violations of the Betweenness Axiom and Nonlinearity in Probability

COLIN F. CAMERER
Graduate School of Business, University of Chicago

TECK-HUA HO*
Department of Decision Sciences, University of Pennsylvania

Abstract

Betweenness is a weakened form of the independence axiom, stating that a probability mixture of two gambles should lie between them in preference. Betweenness is used in many generalizations of expected utility and in applications to game theory and macroeconomics. Experimental violations of betweenness are widespread. We rule out intransitivity as a source of violations and find that violations are less systematic when mixtures are presented in compound form (because the compound lottery reduction axiom fails empirically). We also fit data from nine studies using Gul's disappointment-aversion theory and two variants of EU, which weight separate or cumulative probabilities nonlinearly. The three theories add only one parameter to EU and fit much better.

Key words: expected utility, generalized utility, risk-aversion, prospect theory, Allais paradox, betweenness, disappointment-aversion

Evidence that subjects violate the independence axiom of expected utility theory (EU) has mounted steadily since Allais's (1953) celebrated paradox (see Machina (1987); Weber and Camerer (1987). We describe and dissect empirical violations of a weakened form of independence, called "betweenness." The betweenness axiom states that if X is preferred to Y, then probability mixtures of X and Y must lie between them in preference. (Coupled with continuity, betweenness implies that if X and Y are indifferent, then any mixture of them is indifferent too.) The independence axiom implies betweenness, but betweenness does not imply independence.

There has been a small burst of interest in betweenness for several reasons. One reason is that the axiom seems intuitively appealing; perhaps subjects who violate independence in choice experiments will obey the weaker requirement of betweenness. Another reason is that replacing independence with betweenness has some formal appeal because betweenness amounts to preferences being both (weakly) quasi-convex and

JEL codes: D81, C91
*The financial support of the National Science Foundation (SES 88-09299) is gratefully acknowledged. Comments from Larry Epstein, Rob Porter, John Quiggin, several anonymous referees, and seminar participants at the University of Pennsylvania were helpful.

quasi-concave, two properties that theorists know how to work with. A third reason is that many basic results in economic theory can be derived from betweenness, without using the full force of independence; so betweenness has been widely applied in game theory, auction theory, macroeconomics, and dynamic choice. For example, the crucial choice axiom for proving existence of Nash equilibrium in a finite noncooperative game is betweenness (not independence): Existence requires that if people are indifferent to several pure strategies, they should be indifferent to any probabilistic mixture of those strategies (or "mixed strategy"), a requirement satisfied by betweenness.

Despite recent curiosity about betweenness, there have been relatively few direct tests of it. We conducted some new tests, and surveyed previous studies, to dissect betweenness empirically, much as previous studies dissected independence (and other axioms).

Our study makes three contributions.

First, previous studies tested betweenness applied to single-stage lotteries. These tests make the joint assumption that compound (two-stage) lotteries are reduced to single-stage equivalents, and betweenness applies to compound lotteries. We find that betweenness is, in fact, an adequate descriptive axiom in its compound-lottery form, but not in reduced-lottery form. (The same result appears to hold for independence. Both results add to indirect evidence that the reduction assumption is a surprisingly poor descriptive axiom. See Segal (1990).)

Second, we survey nine studies of betweenness. In most studies, betweenness is systematically violated. The violations are not always as widespread as well-known violations of independence, like the paradoxes of Allais and Kahneman-Tversky. However, in some studies, the violations are very dramatic, and there is generally no doubt that the axiom can be rejected as a descriptive principle (for reduced-form lotteries).

The important question is whether any alternative theory can explain the pattern of betweenness violations. To explain the pattern of violations across studies, we fit single-agent stochastic choice models which assume each subject has the same preference but does not make the same choice every time (he/she chooses one gamble over another with a probability that depends on the difference in their utilities). We estimated parameters for three generalizations of expected utility which are nonlinear in probability, by finding parameter values which maximized the likelihood of observing the choices actually made in the nine studies. Disappointment-aversion and theories in which probabilities or cumulative probabilities were weighted nonlinearly both fit the data better than EU. Each theory we tested adds only one extra parameter to expected utility, and the parameter estimates were remarkably stable across studies. We think any of the theories are therefore parsimonious and precise enough to be useful in economic theorizing. Theorists should no longer continue to use EU just because they think there are no good alternatives.

Our third contribution shows that apparent violations of betweenness are *not* entirely due to violations of transitivity, as has been suggested.

The article has several sections. In the next section (1) we review the basic axioms. Section 2 describes some applications of theories based on betweenness, rather than independence, to game theory, auctions, and macroeconomics. Section 3 describes previous empirical studies of betweenness. Then we test whether apparent betweenness violations might actually be due to violations of transitivity (section 4) or compound lottery reduction (section 5). Section 6 describes maximum-likelihood fits of single-agent stochastic choice models to nine studies. Section 7 concludes.

1. The reduction, independence, and betweenness axioms

We are interested in preferences over single-stage lotteries of "mixtures"—known probability distributions over dollar outcomes. Denote elements of the set of single-stage lotteries by A, B, C, etc. We are also interested in two-stage compound lotteries, lotteries with other lotteries as payoffs. Denote a compound lottery, giving a p chance of A and a $(1 - p)$ chance of B, by $(p, A; 1 - p, B)$. Denote the same lottery, reduced to its single-stage equivalent by multiplying probabilities and then adding probabilities of identical outcomes, by $pA + (1 - p)B$.

We concentrate initially on preferences for single-stage lotteries. Assume preferences are weakly ordered (complete and transitive), and continuous. Formally:

Mixture-weak order: For all A and B, either $A > B$, $B > A$, or $A \sim B$. If $A > B$, $B > C$, then $A > C$.

Mixture-continuity: For all $A > B > C$, there exists a unique $p \in (0, 1)$ such that $pA + (1 - p)C \sim B$.

We can define analogous axioms for preferences over compound lotteries. For example, compound-weak order just substitutes compound lotteries for A, B, and C.

We now define several crucial axioms. Reduction of compound lotteries, or ROCLA, assumes a person is indifferent between a compound lottery and its reduced equivalent. Formally,

Reduction: $(p, A; 1 - p, B) \sim pA + (1 - p)B$.

Note that assuming reduction, along with transitivity, forces the preference orders for single-stage lotteries (mixture-weak order) and compound lotteries (compound-weak order) to be consistent. That is, if $(p, A; 1 - p, B) \sim (p, A; 1 - p, C)$ then $pA + (1 - p)B \sim pA + (1 - p)C$.

Working only with reduced lotteries for the moment, we define independence as follows (see Segal (1990):

Mixture-independence: If $A > B$ then $pA + (1 - p)C > pB + (1 - p)C$ for all $p \in (0, 1]$, for all C.

Independence states that preferences among compound gambles, expressed in their reduced form, should be independent of any common consequence with identical probability (C, in the definition). Since there is substantial evidence against mixture-independence as a descriptive principle (see, e.g., Camerer (1992)), theorists have explored ways to weaken the axiom to explain observed paradoxes. One way to weaken independence is to restrict it to apply only to special sets of lotteries.[1] If we require C to be either A or B—restricting independence to apply only to mixtures of two lotteries A and B with one of themselves (either A or B)—then the betweenness axiom results. Formally,

Mixture-betweenness: If $A > B$ then $A > pA + (1 - p)B > B$ for all $p \in (0, 1)$.

Betweenness requires that every probabilistic mixture of two lotteries A and B should lie between them in preference (hence the term "betweenness"). Adding continuity implies an indifference form of betweenness, in which $A \sim B$ implies that $A \sim pA + (1 - p)B \sim B$. Intuitively, one should be indifferent to randomly mixing equally good outcomes.

One can define independence and betweenness axioms which apply only to comparisons between compound lotteries. Formally,

Compound-independence: If $A > B$ then $(p,A;1 - p,C) > (p, B;1 - p,C)$ for all $p \in (0, 1)$, for all C.
Compound-betweenness: If $A > B$ then $A > (p,A;1 - p, B) > B$ for all $p \in (0, 1)$.

Note that ROCLA and each of the compound-form axioms implies the corresponding reduced-form axiom. For example, if ROCLA holds, then independence of compound lotteries implies that independence holds when applied to the compound lotteries' reduced-form equivalents. Empirical failure of a reduced-form axiom could, therefore, be due to violations of ROCLA, the compound-form axiom, or both (as Luce (1990) and Segal (1990) have argued).

Figure 1 gives a set of gambles (denoted triple (T, U, V)) which help illustrate the axioms. (The gambles depicted were used by Prelec (1990) and in replications reported below.) The numbers on tree branches represent which tickets, drawn from a box containing tickets numbered 1 through 100, select each branch.

The middle gamble Uc is a compound gamble which yields the less risky gamble (T) with probability 16/17, and the more risky gamble (V) with probability 1/17. Compound-betweenness requires $T > Uc > V$, $V > Uc > T$, or $T \sim Uc \sim V$, which are intuitively appealing patterns. The bottom gamble Ur is the compound gamble Uc reduced to a single-stage gamble. Mixture-betweenness requires $T > Ur > V$, $V > Ur > T$, or $T \sim Ur \sim V$, which are much less intuitive: A risk-averse person who prefers T to V, but overvalues the small (1%) chance of getting \$30,000 in Ur, could choose $Ur > T$ and $Ur > V$ and violate betweenness. Note that in the compound form, betweenness requires only a kind of "randomization-neutrality," which is extremely appealing. But when gambles are reduced, the fact that Ur is a mixture of T and V is harder to see. Instead of randomization-neutrality, mixture-betweenness requires something very close to linearity in probability.

Some graphical implications of independence and betweenness can be illustrated in the elegant Marschak-Machina triangle (first used by Marschak (1950); see also Machina (1982); see figure 2.) The triangle shows the set of gambles with three possible outcomes, denoted X_H, X_M, and X_L (ranked in that order). A gamble over these three

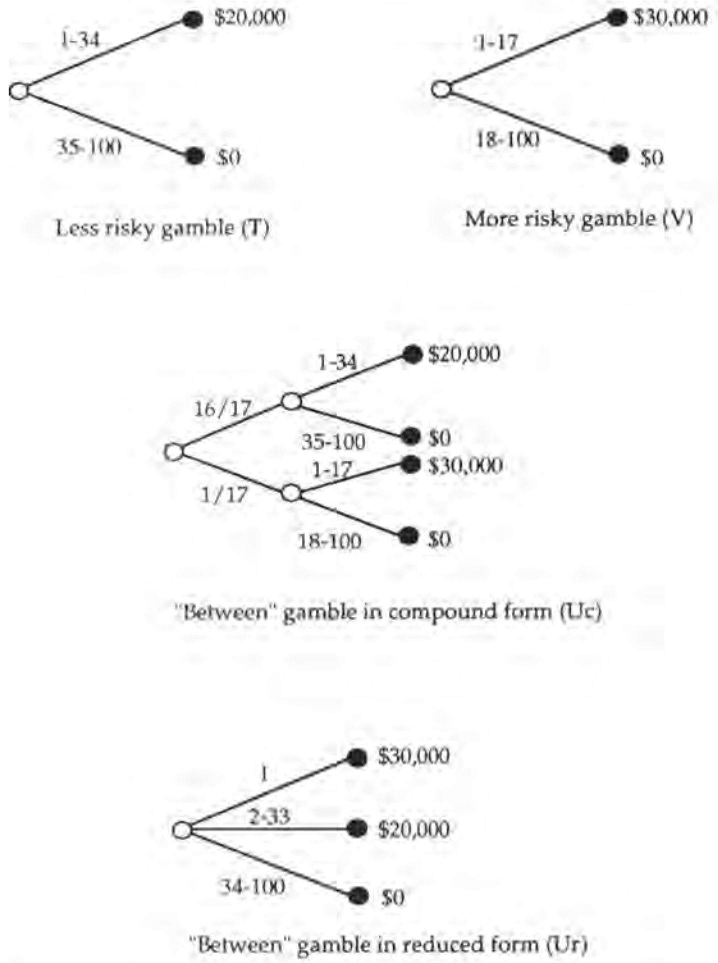

Figure 1. One of the gamble triples used in our experiments

outcomes is a vector of probabilities, p_H, p_M, and p_L. Since $p_M = 1 - p_H - p_L$, each gamble is a point in $p_H - p_L$ space. Indifference curves connect sets of equally preferred gambles.

First note that a reduced-form mixture of D and F, denoted E, is a point on the line connecting D and F. (The compound version of the mixture cannot be shown.) Mixture-betweenness requires that if $D \sim F$, then every mixture of them should be indifferent to D and F. The indifference curve connecting D and F must therefore pass through the set of points which denote the probability mixtures $pD + (1 - p)F$. That set is the line connecting D and F. So sets of equally preferred points—indifference curves—are straight lines.[2]

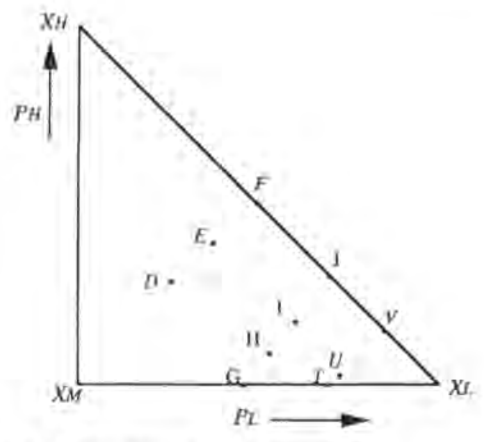

Figure 2. Gambles in Marschak-Machina Triangle

1.1. Betweenness, quasi-concavity, and quasi-convexity

The triangle diagrams can also be used to illustrate ways in which betweenness might be violations. If indifference curves are not linear, they may be bowed in either of two directions, reflecting quasi-convexity or quasi-concavity (or they could have several inflection points).

If either S or R (or both) is (are) preferred to $pR + (1 - p)S$, preferences are quasi-convex. (In general, a function $f(\bullet)$ is strictly quasi-convex if for every $S \neq R$ and $p \epsilon (0,1), f(pS + (1-p)R) < \max(f(S), f(R))$.) The left half of figure 3 shows quasi-convex indifference curves. In the diagram, R is preferred to S (it lies on an indifference curve closer to the northwest, the direction in which preference increases). By betweenness, the mixture $pS + (1-p)R$ should lie on an intermediate indifference curve, but instead it lies on the worst curve of the three. Indifference curves in the triangle diagram are concave, and "worse-than sets" (the sets $\{A{:}B \geqslant A\}$ for each point B) are convex.

The right half of figure 3 shows quasi-concave indifference curves. (A function is strictly quasi-concave if for every $S \neq R$ and $p \epsilon (0, 1), f(pS + (1 - p)R) > \min (f(S), f(R))$.) If S is preferred to R, then the mixture $pS + (1 - p)R$ should lie on an intermediate indifference curve (under betweenness), but under quasi-concavity it lies on a better curve than S and R. Quasi-concavity, therefore, reflects randomization-preference. Indifference curves are convex and "better-than sets" are convex.

There is one subtlety in the way in which we use the terms betweenness, quasi-concavity, and quasi-convexity, and in which we test these properties. If we observe $S > pS + (1-p)R > R$, then betweenness is satisfied for that triple of choices. That pattern *also* satisfies strict quasi-convexity, because $S > pS + (1 - p)R$, and satisfies quasi-concavity because $pS + (1-p)R > R$. For simplicity, we refer to patterns like these as satisfying betweenness (though they satisfy quasi-convexity and quasi-concavity too). If we observe $pS + (1 - p)R > R$ and $pS + (1 - p)R > S$, we call such a pattern "quasi-concave," because it clearly violates betweenness and quasi-convexity, but satisfies quasi-concavity. The same holds true for quasi-convexity.

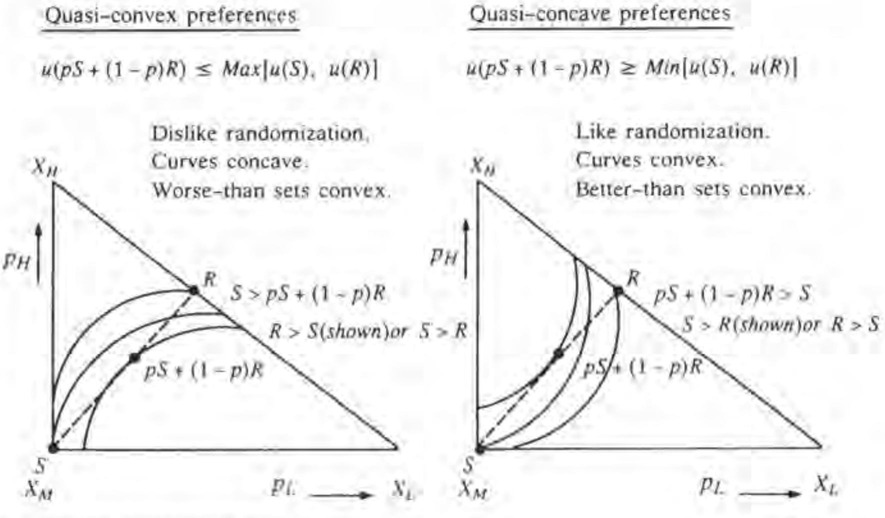

Figure 3. Possible Violations of Betweenness

Our procedure is not the most efficient way to separate betweenness from quasi-convexity and quasi-concavity, because what we call a betweenness-satisfying preference pattern satisfies the other properties too. A smart referee pointed out that the best test to distinguish betweenness from (strict) quasi-convexity and quasi-concavity uses two gambles for which $S \sim R$. Betweenness then requires $pS + (1-p)R \sim S \sim R$ for all p, but strict quasi-concavity requires $pS + (1-p)R > R$ and $pS + (1-p)R > S$, and strict quasi-convexity requires $R > pS + (1-p)R$ and $S > pS + (1-p)R$). This more powerful test requires finding gambles S and R which subjects prefer equally, which takes more experimental time than having subjects choose between a prespecified S and R. As a practical matter, our choice-based tests do the job, because they detect plenty of betweenness violations. But further experiments, with the indifference-based procedure, would produce more evidence of betweenness violation and could be useful in guiding further theory development.

1.2. The intuitive appeal (or lack thereof) of betweenness

Betweenness has been widely applied in recent work, and is thought by many people to be a leading heir to the throne of independence. Yet the intuitive appeal of betweenness is not much greater than for independence.

Some arguments against independence rely on psychological complementarity between lotteries A and C, for example, which makes the mixture $pA + (1-p)C$ more attractive than $pB + (1-p)C$, even though $B \sim A$. These arguments against independence can be turned into arguments in favor of betweenness only if there is no similar complementarity between A and B which might make the mixture $pA + (1 - p)B$ either better than A or worse than B (assuming $A \sim B$). Many of the complementarity arguments against independence hinge on certainty effects—e.g., A is a certain outcome, and hence preferred to the gamble B, but $pA + (1 - p)C$ is uncertain and loses its edge over $pB + (1 - p)C$. Certainty effects of this sort are unlikely in betweenness. A related argument is that B is better than A, but adding a little chance of A is best of all (so $pB + (1 - p)A$ is preferred to B *and* to A). This kind of complementarity will justify betweenness violations, much as certainty effects lead to independence violations (see Prelec, 1990).

Another way to express the intuitions underlying independence and betweenness is suggested by Machina's (1982) "local utility function" approach. He points out that even if a preference order violates the independence axiom in general, comparisons between two gambles that are similar will be well approximated by their expected utility, calculated using a "local utility function" specific to the two gambles. The independence axiom then requires that the local utility functions for A, B, and C all be the same (and since C is arbitrary, that means one utility function must be used). Betweenness, in contrast, requires that the same utility function be used *for all gambles with the same (implicit) expected utility* (Dekel, 1986). This weakening of the single-utility-function property, while ingenious, seems not much more appealing than the single-function property itself.

2. Betweenness in generalized utility, game theory, and macroeconomics

Betweenness is widely used to weaken the independence axiom and develop formal generalizations of EU. In weighted utility theory (Chew and MacCrimmon, 1979; Chew, 1983), a weakened form of independence implies betweenness.[3] (In the triangle, weighted utility indifference curves must be linear and intersect at a point outside the triangle.) *S*kew-*S*ymmetric *B*ilinear utility (SSB) (e.g., Fishburn (1988)) implies betweenness too (it stems from a convexity assumption). Bordley and Hazen (1992) show that SSB and weighted utility can be interpreted as EU with "suspicion" about probabilities of high outcomes; their suspicious-EU theory therefore satisfies betweenness. Implicit EU (Dekel, 1986; cf. Chew (1989) assumes only betweenness, continuity, and weak order. Gul's (1991) disappointment aversion theory assumes betweenness too (cf. Bordley (1992)).

Indeed, the only major generalizations which *do not* assume betweenness are those in quadratic class (Machina, 1982; Chew, Epstein, and Segal, 1991) and those in the rank-dependent or cumulative class (Quiggin, 1982; Yaari, 1987; Luce, 1988; Green and Jullien, 1988; Segal, 1990; Tversky and Kahneman, 1992. See also Chew and Epstein, 1990).

Betweenness has proved useful in applications too. A cornerstone of noncooperative game theory is the proof that Nash equilibria always exist in finite games. The proof requires that players who are indifferent between pure strategies be indifferent between probabilistic mixtures of them, which requires betweenness but not independence. Quasi-concavity is sufficient too (see Debreu (1952)). Crawford (1990) proved the existence of an equilibrium in *beliefs* when preferences are quasi-convex (assuming players can commit to strategies).

Auction experiments have shown that prices in Dutch auctions and first-price auctions are systematically different (e.g., Cox, Smith, and Walker (1983)), though equilibrium bids under EU should be the same. Weber (1982) showed that weighted utility, which obeys betweenness but not independence, could explain the observed difference in prices (see also Chew (1989)). Karni and Safra (1989) showed that even if independence is violated, if bidders obey betweenness, they will bid up to their reservation prices in an ascending-price English auction.

In conventional macroeconomic models of asset pricing, the degree of risk-aversion (or concavity of marginal utility) and the intertemporal substitutability of consumption are confounded, because they are expressed by a single parameter. Epstein and Zin (1989, 1991a) showed that recursive utility (Kreps and Porteus, 1978, 1979) preferences can disentangle the two parameters, but a third variable—preference for resolution of consumption lotteries—is still entangled. Epstein and Zin (1991b) found that a generalization of EU which relies on betweenness instead of independence (due to Gul (1991)) does a much better job explaining aggregate consumption patterns and asset returns than does the recursive utility theory (which assumes independence).

Gul and Lantto (1990) point out three surprising implications of betweenness in dynamic choices. Betweenness implies: A weak kind of "consequentialism," or independence of outcome utilities from unreached outcomes (cf. Machina (1989)); solvability of decision trees using dynamic programming; and expected utility maximization when utilities depend on the choice set.

3. Empirical violations of betweenness

Since betweenness is an appealing replacement for independence, and has proved useful in deriving generalizations of expected utility and applying them to game theory, auction theory, and macroeconomics, empirical testing of it is important.

Several tests have been conducted. The first test was done by Becker, DeGroot and Marschak (1963). Their subjects chose one gamble from the triple, R, S, and $B = .5R + .5S$ (in reduced form). If their preferences satisfy betweenness, subjects should *never* choose the mixture B (unless they are indifferent among R, S, and B), but they actually chose it in 30% of the triples, suggesting strictly quasi-concave preferences.

A similar test was done by two psychologists, Coombs, and Huang (1976). They were motivated by "portfolio theory," the idea that preferences depend on expected value and risk, and preference for risk might be single-peaked. Portfolio theory implies betweenness or quasi-concavity.[4]

Coombs and Huang's results, and the results of several other studies, are summarized in figure 4.[5] For each study, figure 4 shows the approximate locations of the gamble pairs in the Marschak-Machina triangle, and the fraction of subjects obeying betweenness or exhibiting quasi-concavity and quasi-convexity. "On-border" pairs include at least one gamble on a border of the triangle (i.e., at least one of p_H, p_M, or p_L are zero); "off-border" pairs consist of gambles inside the triangle (p_H, p_M, and p_L are strictly positive).

Authors	Location of Gambles in Triangle	Gain			Loss			Mixed-Gain-Loss		
		Betweenness	Quasi-concave	Quasi-convex	Betweenness	Quasi-concave	Quasi-convex	Betweenness	Quasi-concave	Quasi-convex
Coombs and Huang (1976) Experiment I	Cannot graph. L & M are two-outcome gambles; B is a four-outcome gamble							54.4 (N=26)	26.7	18.9
Experiment II	Off-border							86.2 (N=50)	8.5	5.3
Chew and Waller (1986)	On-border	73.2 (N=56)	25.0	1.8	67.9 (N=56)	10.7	21.4	75.9 (N=56)	18.8	5.4
Conlisk (1987) Experiment I	On-border	36.8 (N=152)	9.9	53.3						
Camerer (1989) Hypotenuse pairs	On-border	70.0 (N=258)	7.5	22.5	69.0 (N=253)	23.0	8.0			
Edge pairs	On-border	69.0 (N=215)	26.0	8.5	73.0 (N=214)	9.0	16.0			
Camerer (1992)	Off-border	72.8 (N=82)	15.2	12.0	79.0 (N=86)	12.0	9.0			
Prelec (1990)	On-border	24.0 (N=33)	76.0	0.0	34.2* (N=41)	4.9	60.9			
Gigliotti and Sopher (1993) Treatments 1 & 3	On-border	52.7 (N=281)	7.1	40.2						
Treatment 2	Off-border	69.0 (N=184)	12.0	19.0						
Battalio, Kagel, and Jiranyakul (1990) Set 1	On-border	69.7 (N=36)	7.0	23.3						
Set 2 and 3	On-border				57.0 (N=34)	34.1	8.9			
Bernasconi (1994) Hypotenuse pairs	On-border	49.5 (N=100)	6.5	44.0						
Edge pairs	On-border	51.5 (N=100)	47.5	1.0						

*Data collected by Camerer and Ho using Prelec stimulus (T, U, V) with negative amounts

Figure 4. Violations of betweenness across different studies in gain, loss, and mixed-gain-loss gambles

Several patterns are apparent across studies. Betweenness is violated frequently and systematically in most of the studies. Violations are weaker and more evenly divided in studies using off-border gambles (Coombs and Huang, experiment II; Camerer, 1992; and Gigliotti and Sopher, treatment 2). The difference between on-border studies (higher violation rates) and off-border studies (lower violation rates) suggests that nonlinearity in weighting low probabilities might be an important source of violations (see also Camerer (1992), pp. 227–229).

The pattern of violations in on-border gambles is complex. First consider gambles over gains (the middle column of figure 4). Some studies show quasi-concavity, and some show quasi-convexity. There appears to be no simple way to predict violations from gamble location. For instance, quasi-convexity seems to occur when one gamble is in the lower corner of the triangle (Conlisk, Gigliotti and Sopher, treatments 1 and 3), but Chew and Waller, and Camerer (edge pairs), used a gamble in the lower corner and discovered quasi-con*cav*ity. Furthermore, the patterns for losses are reflections of the patterns for gains (quasi-concavity in one domain implies quasi-convexity in the other, and vice versa). We show in section 6 that there *are* rather simple ways to explain these complicated patterns, by variants of expected utility theory which are nonlinear in the probabilities.

Next we report three experiments which replicate some of the results in figure 4, test whether transitivity violations could explain them, and test compound-betweenness directly.

4. Transitivity and betweenness violations

A test of betweenness uses two gambles, denoted S and R (for *s*afer and *r*iskier), and a mixture between them, denoted B. Three pairs can be constructed from the three gambles, (S, R), (S, B), and (R, B). Subjects usually make choices in two of the three pairs.

The problem with this procedure is that transitivity is used to *infer* the choice in the third pair. A transitivity violation may then be mistaken for a violation of betweenness (see Gigliotti and Sopher (1990), pp. 28–30).[6] For example, suppose a subject chooses $S \succ R$, and $B \succ S$. (We do not observe the choice between B and R.) We infer from $B \succ S$ and $S \succ R$, assuming transitivity, that $B \succ R$. Since the subject picked $B \succ S$ and (by inference) $B \succ R$, she therefore appeared to violate betweenness. But she might have picked $R \succ B$ if she had been asked. Then she would violate transitivity, not betweenness.[7] Alternatively, suppose a person chooses $S \succ R$ and $S \succ B$. Transitivity implies nothing about the choice between B and R. In most studies, it is assumed that $B \succ R$ and that betweenness is satisfied. But if $R \succ B$, then betweenness is actually violated (and transitivity is obeyed).

Put differently, in tests with two pairs, there are four possible patterns of preference. Tests with three pairs have eight patterns. Some of the eight patterns are intransitive, betweenness-obeying patterns that are mistaken for betweenness violations in the two-pair test. And some of the eight patterns are betweenness violations which are invisible in the two-pair test. It is an empirical question whether there will be more total violations of betweenness, or less, in the proper three-pair test, as compared to the limited two-pair test.

Table 1 shows the ten gambles used in our experiments. For gambles D-J, $X_H = \$200$, $X_M = \$80$, $X_L = \$0$. Gambles T-V were the same, except $X_M = \$120$. Figure 2 shows the location of the gambles in a Marschak-Machina triangle.

We conducted experiments using two subject pools. One group was 53 high school juniors attending a summer class at Penn. The second were 72 Wharton MBAs taking a course on negotiation. The subjects made three pairwise choices $((S, R), (S, B)$, and $(B, R))$ in all four gamble triples (high schoolers) or two of the four triples (MBAs).[8] (Their choices express *strict* preference, since indifference was allowed but very few subjects expressed indifference.) The instructions were those from Camerer (1989). After they made choices, a choice by one student in the group was picked at random and the chosen gamble was played.

The two subject pools' choices are significantly different ($\chi^2(31) = 74.1$, $p < .001$); the high school kids were more risk-averse (they chose *SSB* much more often than the MBAs did). But the crucial difference between 2- and 3-pair betweeness violation rates is similar in the two groups, so table 2 reports all the data pooled together.

Table 2 shows the results from the three-pair test, and how the results *would* have looked if only the two choices (S, R) and (S, B) were made. Labels on the left side of the table show which categories (betweenness, quasi-concavity, and quasi-convexity) the two-pair results would fall into. For example, the first two lines of the tables show the number of subjects choosing *SSB* (i.e., $\underline{S} \succ R$, $\underline{S} \succ$ B, and $\underline{B} \succ$ R) and *SSR*. In a two-pair test, these data would be combined in the category *SS* ($\underline{S} \succ$ R, $\underline{S} \succ$ B), which is consistent with betweenness. The labels on the right hand side of the table show which categories (including intransitive) the same results would fall into in a three-pair test. For example, *SSB* would be consistent with betweenness, SSR with quasi-convexity. The bottom of table 2 gives the total number of patterns as classified by each test.

Half the MBA choice patterns and a third of the high school student patterns violate betweeness (46% overall), rates which are comparable to those in earlier studies. For

Table 1. Gambles used in the experiments

Gamble	PL	PM	PH
D	0.3	0.4	0.3
E	0.4	0.2	0.4
F	0.5	0.0	0.5
G	0.4	0.6	0.0
H	0.5	0.4	0.1
I	0.6	0.2	0.2
J	0.7	0.0	0.3
T	0.66	0.34	0.00
U	0.67	0.32	0.01
V	0.83	0.00	0.17

Note: PL, PM, and PH are probability for the worst, middle, and thebest outcomes. Four gamble triples were constructed from these gambles: (D,E,F), (G,H,I), (H,I,J), (T,U,V)

Table 2. Transitivity and betweenness violations (pooled data, $n = 125$)
(S vs R) (S vs B) (B vs R)

2-pair categorization	3-pair pattern	Gamble Pairs				3-pair categorization
		DEF	GHI	HIJ	TUV	
Betweenness	SSB	37	29	33	17	Betweenness
	SSR	9	6	10	0	q-convex
Betweenness	RBR	14	10	8	3	Betweenness
	RBB	1	7	1	4	q-concave
Quasi-concave	SBB	6	21	6	76	q-concave
	SBR	9	8	13	4	Intransitive
Quasi-convex	RSR	6	2	9	1	q-convex
	RSB	4	0	1	1	Intransitive

		DEF	GHI	HIJ	TUV
2-pair test	Betweenness	61	52	52	24
	q-concave	15	29	19	80
	q-convex	10	2	10	2
3-pair test	Betweenness	51	39	41	20
	q-concave	7	28	7	80
	q-convex	15	8	19	1
	Intransitive	13	8	14	5
	n	86	83	81	106

both groups, violations show strong quasi-concavity for triples *GHI* and *TUV* (the latter result replicates Prelec (1990).

The bottom of table 2 shows that the three-pair test has almost exactly the same number of betweenness violations (a total of 165 out of 356) as the two-pair test (which has 167). The hypothesis that many of the patterns which appear quasi-concave or quasi-convex in the two-pair test are actually intransitivities, revealed by the three-pair test, appears to be flat wrong. The three-pair test reclassifies some apparent betweenness violations as intransitivities (patterns *SBR* and *RSB*), but detects new patterns which appeared to be consistent with betweenness but actually violate it (*SSR* and *RBB*), taking away from betweenness as many patterns as it gives back.[9]

5. Compound lottery reduction and betweenness

Previous studies of betweenness, except Conlisk's (1987), actually test *mixture*-betweenness by reducing the compound lottery $B = (p, S; 1 - p, R)$ to its single-stage equivalent before presenting choices to subjects. Recall that mixture-betweenness is implied by the combination of compound-betweenness and reduction (equivalence of compound and reduced gambles). Apparent violations of mixture-betweenness might therefore be due to violation of reduction, violation of compound-betweenness, or both. Since the reduction axiom is often violated in other contexts (Segal, 1990; Luce, 1990), compound-betweenness might well hold, while mixture-betweenness is violated, because reduction is violated.

There are several analogous studies of the mixture-independence and compound-independence axioms. Most find that compound-independence is satisfied, even in the well-known cases where mixture-independence is violated (Luce, 1990).

To test compound- and mixture-betweenness, we ran experiments with four groups of Wharton MBAs ($n = 144$), using the gambles in table 1. Subjects made choices involving a compound representation of the mixture gambles and choices involving a reduced-form representation (as in figure 1). Most of the subjects made two of the three possible choices in each triple.[10] For each group of 36 subjects, a pair was chosen for one subject and the preferred gamble in the pair was played. (Two happy subjects got \$200; one got \$80; and one got \$0.)

Figure 5 summarizes results. The first number in each cell is the number of patterns observed using the reduced-form of the mixture B (denoted Br); the second number is the number of patterns using the compound form (Bc). For example, for the triple (G,H,I) (the upper right of figure 6), 19 people exhibited the quasi-concave pattern $S > R$ and $B > S$ using the reduced-form mixture Br; 10 people exhibited that pattern using the compound-form mixture Bc.

The total number of betweenness violations is roughly the same using Br and Bc (χ^2 (2) = .55, insignificantly different), except for triple (T,U,V). That is, the total of the two off-diagonal cell numbers out of parentheses (Br) and the total in parentheses (Bc) are about the same (e.g., the totals are 20 and 16 for triple (D,E,F)). However, the violations are much less systematic when the compound form Bc is used.[11]

In all four cases the violations are quite lopsided using Br—showing quasi-convexity for triples (D,E,F) and (H,I,J) and quasi-concavity for (G,H,I) and (T,U,V)—and much less lopsided using Bc.[12] Curiously, subjects are also uniformly more risk-averse for choices involving Bc than for choices involving Br.

In a replication of Prelec (1990), using triple (T,U,V), there is overwhelming quasi-concavity using Br (41 of 53 subjects), which virtually disappears using Bc (only 6 of 53). Compound-betweenness is rarely violated while mixture-betweenness is commonly violated. The .01 probability of earning \$30,000 looms large in the reduced-form Br, but in the compound-form Bc, it arises from the joint probability 1/17(.17) and seems less important (see figure 1).

Since mixture-betweenness requires reduction *and* compound-betweenness, the drop in violation rates from mixture-betweenness to compound-betweenness that is apparent in the figure 5 tables is apparently due to violations of reduction. To check directly whether reduction is violated, for each of the four triples subjects were asked to choose between Br and Bc directly. Their choices are given below each matrix in figure 5. Half were allowed to express indifference (denoted "indifference allowed") and half were not

Triple (D,E,F)

(p(0),p(80),p(200))
S (0.3,0.4,0.3)
R (0.5,0.0,0.5)
Br (0.4,0.2,0.4)

	Br (Bc)	R	
S	Betweenness 18 (29)	Quasi-convex 18 (7)	36
R	Quasi-concave 2(9)	Betweenness 29 (22)	31
	20 (38)	47(29)	67

	Br	Bc	I
Indifference allowed	6	20	8
No indifference	9	24	-

Triple (G,H,I)

(p(0),p(80),p(200))
S (0.4,0.6,0.0)
R (0.6,0.2,0.2)
Br (0.5,0.4,0.1)

	S	Br (Bc)	
S	Betweenness 13(22)	Quasi-concave 19(10)	32
R	Quasi-convex 2(12)	Betweenness 32(22)	34
	15(34)	51 (32)	66

	Br	Bc	I
Indifference allowed	15	14	7
No indifference	16	14	-

Triple (H,I,J)

(p(0),p(80),p(200))
S (0.5,0.4,0.1)
R (0.7,0.0,0.3)
Br (0.6,0.2,0.2)

	Br (Bc)	R	
S	Betweenness 15(20)	Quasi-convex 20(15)	35
R	Quasi-concave 2 (9)	Betweenness 27 (20)	29
	17(29)	47(35)	64

	Br	Bc	I
Indifference allowed	9	10	17
No indifference	11	17	-

Triple (T,U,V)

(p(0),p(120),p(200))
S (0.66,0.34,0.00)
R (0.83,0.00,0.17)
Br (0.67,0.32,0.01)

	S	Br (Bc)	
S	Betweenness 7(42)	Quasi-concave 41(6)	48
R	Quasi-convex 2(3)	Betweenness 3(2)	5
	6(6)	32(32)	53
	9(45)	44(8)	

	Br	Bc	I
Indifference allowed	-	-	-
No indifference	35	18	-

Note: 18(29) denotes 18 responses (S > R, B>R) when B is presented in reduced -lottery form (Br) and 29 responses when B is given in compound-lottery form (Bc).

Figure 5. Violations of betweenness in both reduced and compound form gambles

("no indifference"). When given the opportunity, about 30% expressed indifference between *Br* and *Bc,* obeying reduction. Most others chose *Bc* over *Br* for triple (*D,E,F*), and chose *Br* over *Bc* for triple (*T,U,V*). For these two triples, the consistent violations of reduction explain why mixture-betweenness is violated, but compound-betweenness is not. (In triples (*G,H,I*) and (*H,I,J*), the total number of betweenness violations actually *rises* slightly when *Bc* is used, and there is no strong preference for either *Br* or *Bc* in direct comparison.)

Bernasconi (1994) independently studied mixture- and compound-betweenness as we did. In our experiments, presenting mixtures in compound form caused a small reduction in the rate of betweenness violation (except in triple TUV) and a large reduction in the lopsidedness of violations. Bernasconi's data show the opposite: A large reduction in violation rate (from 49% to 32%) and no reduction in lopsidedness. The difference in results is a puzzle.

6. Nonlinearity in probabilities and betweenness violations

We have shown that betweenness violations are frequent and appear systematic within each of several studies, but that the pattern across studies is complex, that apparent betweenness violations in tests with two of three gamble pairs are *not* due to intransitivities which are only visible in tests with all three pairs, and that compound-betweeness is violated less systematically than mixture-betweenness.

The first two findings suggest that nonlinearity in the probabilities might explain the violations. An alternative approach is to generalize betweenness (but one interesting way to do so—the "mixture symmetry" proposed by Chew, Epstein, and Segal (1991)—does not explain the figure 4 patterns.[13])

Nonlinearity can work in many ways. We consider three ways:

1. Gul's (1991) "disappointment-aversion theory" (DAT) is a one-parameter generalization of EU which obeys betweenness.[14] DAT divides gamble outcomes into two sets, depending on whether they are better or worse than the gamble's certainty-equivalent. Formally, denote the (reduced-form) gamble which has a p_i chance of yielding x_i by $(p_1 \ldots p_n, x_1 \ldots x_n)$. Call the certainty-equivalent of a gamble c, and rank the outcomes $x_n \leq x_{n-1} \leq \ldots x_{j+1} \leq c \leq x_j \leq \ldots x_2 \leq x_1$. Call θ the sum of the probabilities of outcomes that produce elation ($\theta = p_1 + p_2 + \ldots p_j$). Define a weighting function for θ, indexed by a free paramenter β, $\omega(\theta) = \theta/(1 + (1 - \theta)\beta)$. (Note that $\omega(\theta)$ is smaller (larger) than θ if β is positive (negative).) Then the modified expected utility under DAT is

$$u(p_1 \ldots p_n, x_1 \ldots x_n)$$
$$= \omega(\theta) \sum_{i=1}^{j} (p_i/\theta) u(x_i) + (1 - \omega(\theta)) \sum_{i=j+1}^{n} (p_i/(1 - \theta)) u(x_i). \quad (1)$$

Intuitively, DAT breaks a gamble's expected utility into two parts, an elation part (the first term) and a disappointment part (the second term). Each part is a partial expected utility (with normalized probabilities). The weights $\omega(\theta)$ and $1 - \omega(\theta)$ reflect the importance of disappointment or elation in the sum (1).[15] A person who is averse to disappointment will overweigh disappointing outcomes $(x_n, \ldots, x_{j+1})$ and underweigh elating ones (much as concave utility "overweighs" low outcomes, expressing risk-aversion). The single parameter β completely expresses disappointment-aversion: If $\beta > 0$ ($\beta < 0$), a person is disappointment-averse (-preferring). Note that if $\beta = 0$, then $\omega(\theta) = \theta$, and DAT reduces to EU.

Gul's disappointment theory can be neatly illustrated in the Marschak-Machina triangle. Figure 6a shows indifference curves under expected utility (which are parallel straight lines). Figure 6b shows indifference curves under Chew-MacCrimmon weighted

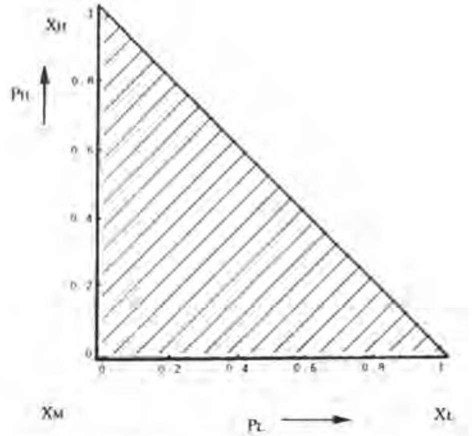

Figure 6a. The indifference map predicted by expected utility theory

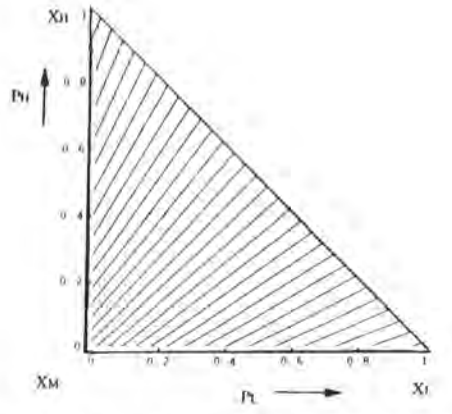

Figure 6b. The indifference map predicted by weighted expected utility theory ($w(X_m) = 0.5$)

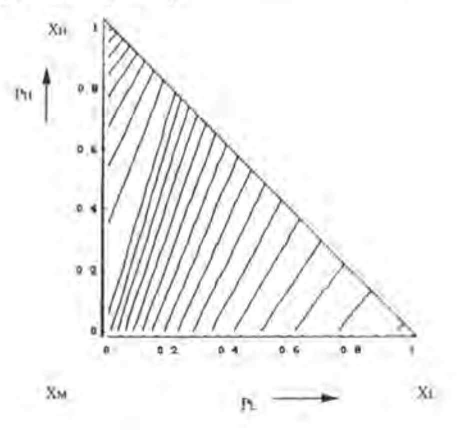

Figure 6c. The indifference map predicted by disappointment aversion theory ($\beta = 2.5$)

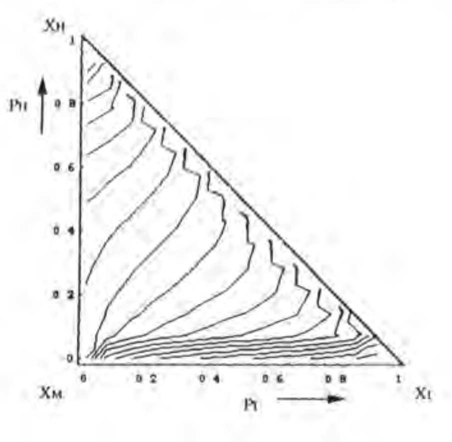

Figure 6d. The indifference map predicted by separable prospect theory ($\gamma = 0.52$)

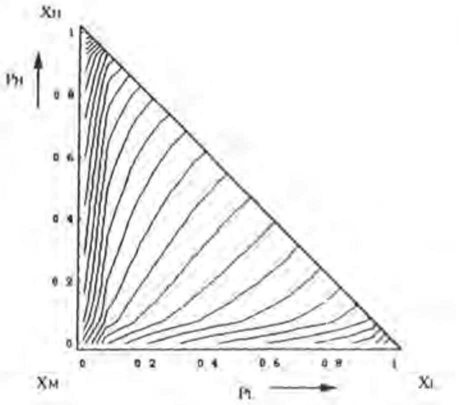

Figure 6e. The indifference map predicted by cumulative prospect theory ($\gamma = 0.56$)

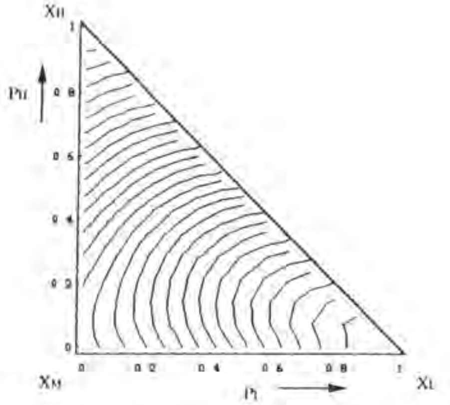

Figure 6f. The indifference map predicted by separable prospect theory ($\gamma = 1.5$)

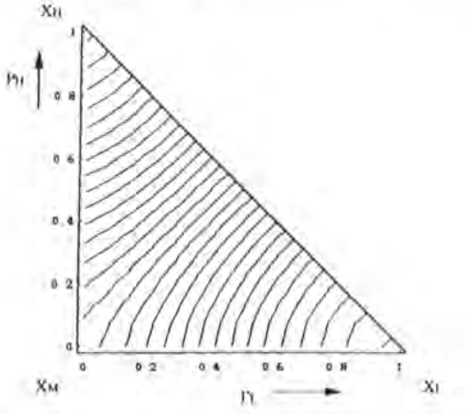

Figure 6g. The indifference map predicted by cumulative prospect theory ($\gamma = 1.5$)

utility (which satisfies betweenness); the curves are straight lines that "fan out," meeting at a point to the southwest. (Curves could also meet at a point to the northeast and "fan in.")

Figure 6c shows indifference curves under disappointment-aversion theory. Curves in the northwest fan in, meeting at some point outside the triangle to the far northeast (not shown in figure 6b). Curves in the southeast fan out, meeting at some point to the far southwest. A middle indifference curve is a "dividing line" that intersects both meeting points. The dividing line could also lie outside the triangle, to the southeast or northwest, allowing uniform fanning in or fanning out (as in figure 6b).

2. A second possibility is that outcome probabilities can be weighted separately and nonlinearly (as in Handa (1977), Karmarkar (1978); Kahneman and Tversky (1979); Viscusi (1989); and others). We consider the functional form

$$u(p_1 \ldots p_n, x_1 \ldots x_n) = \sum_{i=1}^{n} w(p_i)u(x_i) \tag{2}$$

We refer to (2) as "separable prospect theory" (PT), but the estimates we made actually fit a class of nonlinear weighting theories which is more general than prospect theory (as stated in Kahneman and Tversky (1979)).[16]

3. A third approach is that cumulative probability distributions of outcomes can be weighted, then differentials of the transformed cdf used to weight the values of different outcomes. (Or outcomes can be weighted by differentials of the transformed *de*cumulative distribution, one minus the cumulative distribution.) Such "rank-dependent" theories include Quiggin (1982), Yaari (1987), Green and Jullien (1988), Segal (1989), Tversky and Kahneman (1992), Luce (1991), and Luce and Fishburn (1991). To see how rank-dependent weighting works, rank the outcomes x_i with x_1 the best outcome ($x_n \leq x_{n-1} \leq \ldots \leq x_1$). Suppose the *de*cumulative distribution (dcdf) is transformed. Then the gamble $(p_1 \ldots p_n, x_1 \ldots x_n)$ with $p_0 = 0$ by convention) has rank-dependent weighted value:

$$u(p_1 \ldots p_n, x_1 \ldots x_n) = \sum_{i=1}^{n} \left[w\left(\sum_{j=0}^{i} p_j\right) - w\left(\sum_{j=0}^{i-1} p_j\right) \right] u(x_i) \tag{3}$$

We refer to (3) as "cumulative prospect theory" (CPT) although, as with (2), we actually test a more general form than in Tversky and Kahneman (1992). (Also, the analysis below uses only studies with gain gambles, where rank-dependent theories and CPT coincide.)

General properties of theories with separable and rank-dependent probability weights can be tested in many ways (see Camerer (1989, 1992)). For our test, we need a parsimonious parametric specification of w(p) for precision. We studied an elegant one-parameter functional form for probability weights introduced in Tversky and Kahneman (1992):[17]

$$w(p) = p^\gamma/(p^\gamma + (1-p)^\gamma)^{1/\gamma} \tag{4}$$

In "cumulative prospect theory," Tversky and Kahneman apply this form of w(p) to transform the dcdf of a gamble. Cumulative prospect theory is the same as the most general rank-dependent theories except that the value function is more restricted (we consider a more general form), and, as in Luce (1991) and in Luce and Fishburn (1991), different weights are used for gains and losses.[18] Notice that $\gamma = 1$ implies $w(p) = p$, corresponding to linear probability weighting (expected utility).

Other functional forms for $w(p)$ could be explored. For example, a concave (convex) weighting function $w(p)$ will create quasi-concave (-convex) preferences (Röell, 1987); furthermore, global risk-aversion (risk-preference) holds if and only if the utility function is concave (convex) and $w(p)$ is convex (concave) (Chew, Karni, and Safra, 1987). We felt that strictly concave or convex $w(p)$, like the parsimonious power function $w(p) = p^c$, would not fit well, because the data in figure 4 show a clear mixture of quasi-concave and quasi-convex patterns.

Indifference curves from separable *PT* and *CPT* are shown in figure 6d (PT) and figure 6e (*CPT*) assuming specific values of γ—and hence, particular weighting functions w(p) according to (4)—which are estimated from the data (described below). The separable

PT curves are highly nonlinear, violating betweenness, and even slope negatively near the hypotenuse (violating stochastic dominance.[19]) The CPT curves are less dramatically nonlinear and violate betweenness too (but obey dominance). Notice that both sets of curves are sometimes convex, expressing quasi-concavity (e.g., in the southeast) and sometimes concave, expressing quasi-convexity (e.g., toward the northwest). In some cases a single curve is both concave and convex in different parts.

Notice that all three functional forms, DAT (1), PT (2), and CPT (3), exhibit different brands of nonlinearity in probabilities. None of the theories forces $u(B) = pu(S) + (1 - p)u(R)$, when $B = pS + (1 - p)R$ (as EU does, which implies *parallel* linear indifference curves). However, since DAT obeys betweenness, it implies that indifference curves are linear, which means that DAT is linear in probability *along an indifference curve* (i.e., $u(pS + (1 - p)R) = pu(S) + (1 - p)u(R)$ when $u(S) = (R)$).

6.1. *Maximum-likelihood estimation of stochastic choice models*

There are several methods to test whether different theories can explain the diverse results in figure 4. One method allows parameters (and hence, patterns of preference) to vary across subjects and estimates each subject's parameter value, or estimates the percentage of subjects with different parameter values. However, there is no value of β in DAT which can explain the quasi-concave patterns $B > S > R$ and $B > R > S$, or the quasi-convex patterns $R > S > B$ and $S > R > B$, while certain values of γ in PT and CPT *can* explain those patterns. Therefore, a test which allowed subjects' parameter values to vary would likely favor PT or CPT over DAT, because a healthy percentage of subjects violate betweenness in one way or the other in each figure 4 study. DAT cannot explain those subjects; PT and CPT might. Harless and Camerer (in press) report such tests. Indeed, theories based on betweenness (including DAT) are more strongly rejected than theories like PT and CPT.

Here we use an entirely different method, a "single-agent stochastic choice" method. The method has two distinct features.

First, we assume a single pattern of preference for all agents. Of course, we do not think all agents actually have the same preference. But the data shown in our figure 4 survey do not permit us to do much else. We cannot reliably estimate the fraction of subjects with each of several different preference patterns, since in many studies each subject made only a small number of choices (perhaps only two, or one). Therefore, we are forced for the sake of parsimony to assume a single preference.

Another justification for our single-preference approach is that many economic theories assume a single "representative agent," for analytical tractability. As an empirical input to such theorizing, it is useful to know which *single* parameter value and functional form fits best.

Second, since we assume each subject has the same preference, a stochastic element is necessary to explain why subjects' choices would vary if they all prefer the same thing. We used a stochastic choice (or "random utility") model in which the probability of choosing X over Y depends on the difference in their utilities. One can think of the utilities of X and Y as means; the utilities that determine actual choices add a random element to the means, perhaps reflecting further deliberation or pure trembles.[20]

There is much to criticize about the single-agent stochastic choice approach. But it is a simple method for integrating data from many studies (where estimates of individual differences may be impossible or have low power). It also answers the question, "What's the best *empirical* representative agent?" and—most importantly—it complements methods we and others have used which *do* allow individual variation (e.g., Harless and Camerer (in press); Camerer (1992); Kahneman and Tversky (1992); Hey and Orme (in press)).

We assumed a logit form stochastic choice model (following Luce and Suppes (1965), especially p. 335), $P(X > Y) = 1/(1 + e^{u(Y)-u(X)})$. In a stochastic choice model a common estimate of γ or β for all subjects can create heterogeneous choices, because choice percentages naturally vary, according to the strength of preference $u(X) - u(Y)$. (That is, even if $X > Y$ and $u(X) > u(Y)$, Y will be chosen some fraction of the time.) The logit form is especially appealing, because $u(Y) = u(X)$ implies $P(X > Y) = .5$, and *no* free parameters need to be estimated (or even can be).[21]

We tried several specifications for utility functions, including exponential ($u(x) = 1 - e^{\alpha x}$), log-quadratic ($u(x) = \log(x + 1) - \alpha(\log(x + 1))^2$), and power ($u(x) = x^\alpha$). Fits were not identical, but the utility functions ranked competing theories in roughly the same way.[22] Since power utility functions are commonly used, and yielded the best fits for all the theories studied except EU, we report only results with power utility.

An example illustrates how the maximum-likelihood estimation worked, for cumulative (or rank-dependent) theories. In the Battalio, Kagel, and Jiranyakul (1990) study, the gambles used were $S = (.50, \$18; .50, \$27)$, $R = (.90, \$27; .10, \$0)$, and $.4S + .6R = (.20, \$18; .74, \$27; .06, \$0)$. (The latter mixture, denoted B, is shown in reduced form.) Under the rank-dependent theories expressed by equation (3) (setting $u(\$0) = 0$ for brevity),

$$\begin{aligned} u(S) &= w(.5)\,u(\$27) + (1 - w(.5))\,u(\$18) \\ u(R) &= w(.9)\,u(\$27) \\ u(B = .4S + .6R) &= w(.74)\,u(\$27) + (w(.94) - w(.74))\,u(\$18). \end{aligned} \tag{5}$$

Note first that if EU holds, then $w(p) = p$ and $u(B) = .4u(S) + .6u(R)$. Nonlinear $w(p)$, however, can make $u(B) < .4u(S) + .6u(R)$ and can then explain the large fraction of quasi-convex preferences.

In their experiment, subjects chose between S and R, and between B and R. Twenty-three of 36 subjects (69.7%) obeyed betweenness. Ten subjects (23.3%) were quasi-convex, choosing $S > R$ and $R > B$, and three (7%) were quasi-concave ($R > S$ and $B > R$). The maximum-likelihood procedure chooses values of α and γ (using a grid search), then calculates utilities for S, R, and B by plugging $w(p,\gamma)$ from equation (4) and $u(x) = x^\alpha$ into the rank-dependent utility expressions in (5). (For example, $u(R,\gamma,\alpha) = [.9^\gamma/(.9^\gamma + .1^\gamma)^{1/\gamma}](27^\alpha)$.) If we assume that choices in the two pairs are statistically independent,[23] the likelihood function for a particular γ and α, $L(\gamma,\alpha)$, is

$$\begin{aligned} L(\gamma,\alpha) = &\{P(S > R)P(B > R) + P(R > S)P(R > B)\}^{23} \\ &\{P(S > R)P(R > B)\}^{10}\{P(R > S)P(B > R\}^{3} \end{aligned} \tag{6}$$

The first term represents the probability that a subject obeys betweenness (by choosing $S > R$ and $B > R$ or $R > S$ and $R > B$); the second and third terms give the probability of quasi-convexity and quasi-concavity.

Grid search over suitable parameter values of γ (from .27 to 2)[24] and α (between 0 and 2) yields maximum-likelihood estimates of (6). (For DAT, we allowed β to range from -1 to 50.) We estimated parameters in each study separately, using only gain-gamble data, and in all the studies together.

Table 3 shows log likelihoods and maximum-likelihood estimates of the parameter β in disappointment-aversion theory (DAT), and the weighting parameter γ for both prospect theory (PT) and cumulative (rank-dependent) prospect theory (CPT), for each of the studies given in figure 4. The maximum-likelihood estimates, when parameters are restricted to be constant across studies, are shown at the bottom of table 3, both unweighted (each study counts equally) and weighted (by each study's sample size).

Log likelihoods from EU, given in the second column of table 3, provide a benchmark for the other theories. Disappointment (DAT) and the two versions of prospect theory each add one parameter to EU. We can test whether their additional parameters help explain the data with a chi-squared test.[25]

The chi-squared test rejects the EU restriction that $\beta = 0$ in favor of DAT in six of eight studies. Most estimates of β are positive, roughly between 1 and 10. The estimates vary substantially (though in DAT, large variation in β values does not imply large variation in choice patterns). A test of whether the true β is the same in all studies strongly rejects the hypothesis of identical β's ($\chi^2(10) = 102.8$, $p < .001$). Restricting β to be constant across studies, we estimate β to be around 3 (with $\alpha = .225$) and strongly reject EU in favor of DAT.

The EU restriction that $\gamma = 1$ is rejected in favor of either PT or CPT in six and seven of eight studies, respectively, and is strongly rejected across studies. The estimates of γ vary across studies but are generally less than 1. (A test of whether γ is equal in all studies, for each theory, strongly rejects equality with $x^2(10) = 100.8$ and $x^2(10) = 118.2$, both $p < .001$.) The pooled estimates γ are .52 and .56 (and $\alpha = .225$). These estimates are remarkably close to the estimate of .61 for CPT, derived by Tversky and Kahneman (1992), who used a very different procedure (minimizing squared deviations from certainty-equivalents). Wu ((1993), especially p. 18) also inferred a similar weighting pattern from data that test cancellation axioms.

Figure 7 shows a plot of the probability weighting functions for our CPT estimate of $\gamma = .56$. Low probabilities are overweighted and higher ones are underweighted, with a crossover point ($w(p) = p$ around .3. For comparison, figure 7 also shows $w(p)$ with $\gamma = 1.5$, in which the opposite pattern occurs—low (high) probabilities are underweighted (overweighed). Figures 6f and 6g show indifference curves for separable PT and CPT with $\gamma = 1.5$. The curves vary less than with lower values of γ around .5 (cf. figures 6d–6e) because the weighting function with $\gamma = 1.5$ is closer to $w(p) = p$.

6.2. *Two puzzles*

The results of our single-agent, data-fitting exercise have two curious features. The first curious feature is that disappointment-aversion theory assumes betweenness, which is

Table 3. Log likelihoods and parameter estimates for EU, DAT, PT, and CPT (gains gambles only)

	EU	Disappointment Aversion Theory (DAT)		Separable Form Prospect Theory (PT)		Cumulative Prospect Theory (CPT)	
Authors, Theory	Likelihood	$\hat{\beta}$	Likelihood x^2 statistic	$\hat{\gamma}$	Likelihood x^2 statistic	$\hat{\gamma}$	Likelihood x^2 statistic
Chew and Waller (1986)	− 137.5	1.0	− 111.9 51.2***	0.63	− 115.1 44.8***	0.74	− 115.1 44.8***
Conlisk (1987) Experiment 1	− 164.2	7.9	− 152.7 23.0***	1.87	− 142.4 43.6***	0.56	− 142.5 43.4***
Camerer (1989) Hypotenuse pairs	− 242.1	1.3	− 240.5 3.2	1.01	− 241.2 1.8	0.64	− 241.1 2.0
Edge pairs	− 366.6	2.2	− 360.6 12.0***	0.90	− 357.2 18.8***	0.82	− 359.8 13.6***
Camerer (1992)	− 277.0	− 0.2	− 276.8 0.4	1.03	− 276.7 0.6	0.97	− 276.9 0.2
Prelec (1990)	− 25.6	− 0.8	− 25.0 1.2	0.88	− 18.2 14.8***	0.69	− 18.2 14.88***
Gigliotti and Sopher (1993) Treatments 1 & 3	− 286.5	48.9	− 265.7 41.6***	1.51	− 250.6 71.8***	1.87	− 250.6 71.8***
Treatment 2	− 154.0	22.7	− 152.0 4.0*	1.14	− 151.9 4.2*	0.32	− 152.1 3.8*
Battalio, Kagel, and Jiranyakul (1990) Set 1	− 32.3	5.2	− 28.0 8.6**	1.60	− 31.9 0.8	0.72	− 28.0 8.6**
Bernasconi (1994) Hypotenuse pairs	− 202.2	8.7	− 174.1 65.1***	2.00	− 194.9 14.6***	0.28	− 185.3 33.8***
Edge pairs	− 199.9	5.3	− 178.1 43.6***	0.47	− 182.4 35.0***	0.48	− 194.6 10.6***
TOTAL Weighted	− 2163.2	2.5	− 2068.2	0.52	− 2063.3	0.56	− 2082.4
Unweighted	− 1044.8	3.2	− 1002.8	0.52	− 986.4	0.55	− 1005.9

*p < 0.05; ** p < 0.01; *** p < 0.001

clearly violated in the studies summarized by figure 4, but which fits better than EU, and about as well as the theories which allow betweenness violations. The surprisingly good fit of DAT is due to a property of "representative agent" models in general: If there is a mixture of preferences in the population—most subjects obey betweenness, but a minority do not—then a model like DAT, which fits the betweenness-violating minority poorly, could still fit the aggregate data generated by the mixed population well.[26]

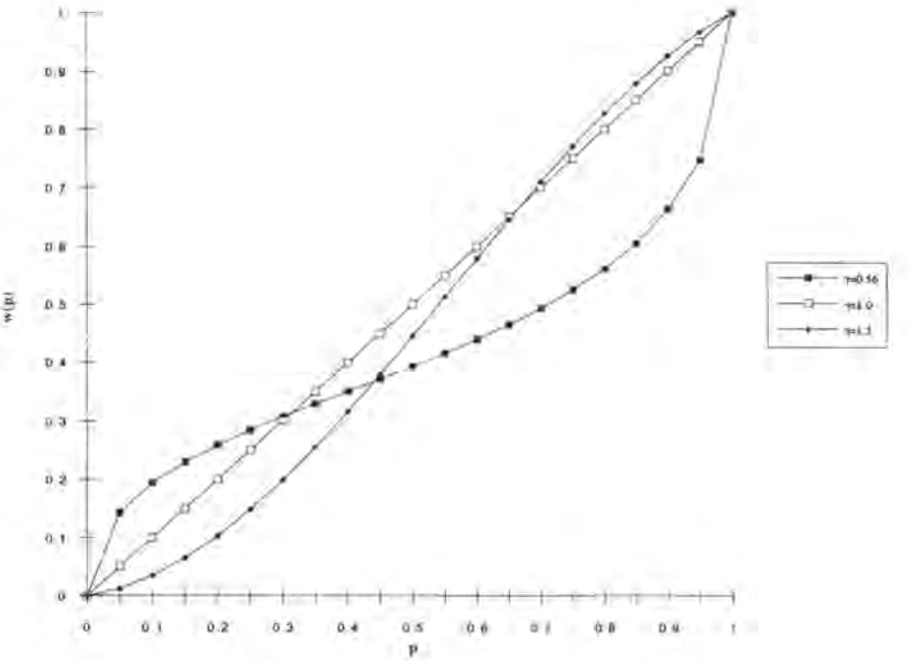

Figure 7. Plots of $w(p)$ for $\gamma = 0.56, 1.5$

The second curious feature of the data is that, overall, separable prospect theory fits the data somewhat better than cumulative prospect theory.[27] (The two fit about equally well in each study). Whether outcome weights are separable or rank-dependent is an important empirical question. More direct tests, estimating weights both ways or exploiting the fact that separable weights imply dominance violations, would be useful (see Wu (1993)).

7. Conclusions

Betweenness is a weakened form of the independence axiom, which states that a probability mixture of two gambles should be between them in preference. Betweenness figures prominently in generalizations of expected utility and their applications to game theory and macroeconomics.

A survey of previous empirical work on betweenness suggested that violations of the axiom are widespread and systematic. We tested whether apparent violations might be due to intransitivity. They were not. We also tested betweenness in its compound form, since previous tests generally used mixtures of gambles reduced to single-stage equivalents. When gambles were presented in compound form, violations of betweenness were somewhat fewer and much less systematic than when gambles were in reduced form. (Similar results have been found for the independence axiom.) The results suggest that theories which assume betweenness (on reduced-form gambles) are skating on thin empirical ice.

Compelling evidence of independence violations inspired lots of interesting theorizing about non-EU alternatives. In the same spirit, we think data on betweenness violations are sufficiently clear that descriptive modelling efforts should turn sharply away from betweenness-based theories and toward theories in other classes, especially those with nonlinear probability weights.

When people make choices among compound gambles, however, both independence and betweenness are more likely to be satisfied. The striking difference between systematic violation of reduced-form betweenness and unsystematic violation of compound betweenness suggests that the way in which gambles are described to people, or "framed" by them, is an extremely important determinant of their choices. For predictive purposes, knowing whether people imagine choices as compound gambles or not is essential to guessing whether they will obey EU or not.

The reduced-form betweenness violations we survey and report seem to be due to nonlinearity in the probabilities. We used eight studies to estimate the best-fitting functional form of nonlinearity in probability, using three specific functional forms.

A theory which relaxes independence but obeys betweenness, disappointment-aversion theory (Gul, 1991), fit the data much better than EU. In two other non-EU forms that we tested, outcome probabilities are weighted separably (and nonlinearly), or outcomes are weighted by differences in cumulative probabilities which are transformed nonlinearly. Estimates of the probability weighting function were remarkably close to the pattern observed by others using very different methods (Tversky and Kahneman, 1992): Low probabilities (below .30) are overweighted and higher probabilities are underweighted.

The superior fit of disappointment-aversion theory and the similarity of the probability weighting estimates across eight studies suggest two brands of parsimonious (one-parameter), reliable empirical replacements for EU. We think it is high time that theorists and others who use expected utility theory as a descriptive theory, should apply some of these functional forms—which add just one parameter to EU—and see if other kinds of anomalies can be explained by using the simple new forms instead of using EU.

Notes

1. For example, Schmeidler (1989) weakened Savage's sure-thing principle—the equivalent of independence in a subjective expected utility approach—by restricting it to apply only to pairs of "comonotonic" acts which rank states in the same way.
2. Independence is stronger. It requires that the mixtures $pD + (1-p)Q$ (where Q is an arbitrary gamble, not shown in figure 2) and $pF + (1-p)Q$ be on the same indifference curve, if D and F are on the same curve. Graphically, independence therefore implies that curves are *parallel* straight lines.

3. Weak independence states: If $A \sim B$, for any p in (0,1) and any C, there is a $q(p)$ in (0,1) such that $pA+(1-p)C \sim q(p)B+(1-q(p))C$. Fix $C=B$. Then the axiom states that $pA+(1-p)B \sim B$ for all p.
4. A single-peaked risk preference could violate betweenness, because a person might prefer a low-risk gamble S to a high-risk gamble R, but like the mixture $pS+(1-p)R$ better than either S or R, because the mixture has the ideal level of intermediate risk. (This quasi-concave pattern is called a "folded ordering".)
5. Harless (1992) conducted an additional study which we learned about too late to include. His paper replicates Prelec (1990) and also finds substantial violations of betweenness (40% quasi-concave, 0% quasi-convex, $n=43$) when Prelec's gambles are slid toward the lower-left corner. He notes that these patterns violate prospective reference theory (Viscusi, 1989) and some betweenness-based theories.
6. We first heard this possibility suggested by Mark Machina at the FUR V conference in Duke.
7. Absent transitivity, a person violates betweenness if $B > R$ and $B > S$ (she prefers the mixture to both its outcomes), or $R > B$ and $S > B$ (she prefers both outcomes to the mixture of them). Note that this definition implies a proper experiment *can* be done with two choices—asking subjects about B versus S, and B versus R.
8. The MBA students also made some choices involving compound versions of the lottery B, which are reported below.
9. Independently of us, Bernasconi (1994) conducted a nearly identical test with similar results: His two-pair test gave 49% betweenness violations, and his three-pair test gave 62% violations. Only 12% of the patterns which are classified as betweenness violations in the two-pair test are due to intransitivity.
10. Some subjects did make all three choices, (S,R), (B,S), and (B,R). Their patterns are quite similar to those in figure 5, because intransitivities visible in the three-pair test are roughly offset by betweenness violations that are not visible in the two-pair test.
11. Systematicity of violations is important, because unsystematic violations could be due to random errors in responses, but systematic violations (by definition) are probably not. (As a result, most experimental studies of EU and its variants measure the degree of violation both ways, by their overall rate *and* systematicity.) However, it is curious that the *overall* violation rates are not reduced when Bc is used instead of Br.
12. The chi-squared statistics testing whether the number of quasi-convex and quasi-concave patterns is independent of whether Br or Bc was used are 8.67, 9.72, 5.09, and 6.98 (all are significant at $p < .025$).
13. When $S \sim R$, mixture symmetry allows a mixture $pS+(1-p)R$ to be preferred to both S and R, or both S and R to be preferred to the mixture (violating betweenness), but requires $pS+(1-p)R \sim (1-p)S+pR$. Mixture symmetry allows preferences to switch from quasi-convex to quasi-concave, as gambles become better and better (in the sense of stochastic dominance); global concavity or convexity are allowed too. The figure 4 data cast doubt on this axiom because there is no tendency to switch from quasi-convexity to quasi-concavity for movements to the northwest corner of the triangle. Bernasconi (1994) tested mixture symmetry most directly. He found that $pS+(1-p)R$ is often preferred to both S and R (quasi-concavity), while S and R are both preferred to $(1-p)S+pR$ (quasi-convexity), implicitly violating $pS+(1-p)R \sim (1-p)S+pR$. (The same pattern is apparent in the hypotenuse and edge pairs in Camerer (1989) (see figure 4) and in indifference curves from PT and CPT in figures 6d–e.)
14. We chose to estimate parameters for DAT, because it has only one free parameter. We excluded some other theories (reviewed in section 3 above). Weighted utility and *SSB* have functions for which no simple specifications have been proposed. Implicit EU with power utility can be estimated by letting each gamble have a different risk-aversion coefficient α (and checking to be sure the implied indifference curves do not violate dominance by sloping downward or violate transitivity by crossing). To fit implicit EU, we fit power utility functions with different α's for each gamble, for studies with only three gambles. The maximum likelihoods were: Conlisk (-143.2), Prelec (-18.2), Gigliotti and Sopher (-251.0 and -152.1), and Battalio et al. (-28.0). The EU restriction (all the α's are the same) is rejected in four of five studies. Implicit EU outpredicts DAT in the first three studies, but it *never* outpredicts cumulative prospect theory (even though it has one *more* degree of freedom).
15. Note that the modified utility in (1) is implicitly defined, because $U(p_1 \ldots p_n, x_1 \ldots x_n)$ determines the certainty-equivalent c, which determines, in turn, whether outcomes are elation outcomes or disappointment outcomes, which affects $u(p_1 \ldots p_n, x_1 \ldots x_n)$. Gul (1991) gives an efficient algorithm for computing utilities using (1) for a given β.
16. Prospect theory assumes several editing rules, uses a different form than (2) when gamble outcomes are strictly positive or negative ("irregular prospects"), and assumes concave value for gains ($\alpha \leq 1$ for power utility). We assumed none of these features in our estimation.

17. Others have proposed similar forms for $w(p)$. Karmarkar (1978) suggested $w(p) = p^\gamma/(p^\gamma + (1-p)^\gamma)$ (which satisfies $w(.5) = .5$, as in Quiggin (1982)). Lattimore, Baker, and Witte (1992) proposed $w(p_i) = \alpha p_i^\beta/(\alpha p_i^\beta + \Sigma p_k^\beta)$. We explored (4), because it has one parameter and we were skeptical that $w(.5) = .5$. We also tested a two-parameter cubic form of $w(p)$ with $w(0) = 0$ and $w(1) = 1$. The cubic form fit no better than equation (4), so we quit exploring it. However, note that the form in (4) uses only one parameter to express two features of $w(p)$, the crossover point at which $w(p) = p$ and the curvature. Trying to capture both features with a single parameter sometimes leads to poor fits. John Quiggin pointed out that the specification $w(p) = p^\gamma/(p^\gamma + (1-p)^\gamma)^\alpha$ includes Karmarkar's form ($\alpha = 1$) and Tversky and Kahneman's form ($\alpha = 1/\gamma$) as special cases; tests of the two-parameter form could then test whether the restrictions $\alpha = 1$ and $\alpha = 1/\gamma$ are satisfied or not.
18. The switch in direction of betweenness violations in gain and loss gambles that is evident in figure 4 supports the assumption that weights differ for gains and losses. Tests like those reported later in this section showed that the reflection assumption, $w_+(p) = 1 - w_-(1-p)$ (where $w_+(p)$ denotes gain weight, and $w_-(p)$ loss weight) fits the figure 4 gain-loss data consistently, but only slightly better, than the equality assumption $w_+(p) = w_-$. So tests and applications of rank-dependent theories should be sensitive to possible differences in weights for gains and losses.
19. The fact that nonlinear, separable weights lead to violations of dominance has been known for a long time (Kahneman and Tversky (1979) include a step in their "editing phase" for detection of dominance violations, and show in their 1986 article that violations are in fact observed when the dominance relation is not transparent.) Intuitively, dominance is violated here because moving away from the hypotenuse (where $p_M = 0$) to the interior triangle raises p_M slightly above zero, overweighting it substantially and making people willing to accept dominated gambles with higher p_L and lower p_H in exchange for the overweighted increment in p_M.
20. Some studies have established the degree of randomness in choice empirically, by asking subjects to choose between the same pair of gambles twice (or more). Subjects reverse preference 15-30% of the time, depending of course on the gamble pair (e.g., Camerer (1989); Wu (1993)). Machina (1985) and Crawford (1988) point out that choices like this are consistent with deterministic preferences that are quasi-concave (but as we show above, global quasi-concavity appears to be rejected by the data in figure 4).
21. If $e^{u(Y)-u(X)}$ is multiplied by a constant, then the property that $u(Y) = u(X)$ implies $p(X \succ Y) = .5$ is violated. Multiplying $u(Y) - u(X)$ by a constant is pointless, since utilities are only determined up to an arbitrary positive multiple. Therefore, no free parameters *can* be added.
22. For example, the log maximum likelihoods across all the studies in figure 4 (weighted) for exponential, log-quadratic, and power utility were -2203, -2121, -2163 (EU), -2170, -2100, -2068 (disappointment-aversion theory, DAT); -2159, -2085, -2063 (separable prospect theory, PT); and -2159, -2098, -2082 (cumulative prospect theory, CPT). Note that the rank of the theories is EU < DAT < PT for all three utility functions. CPT ranks below DAT in power utility and above it in the other two specifications. We report only results with power utility functions below, because it fits best (except for EU), and the estimates of α were most stable across theories using power utility.
23. Statistical independence means that $P(S \succ R \mid .4S + 6R \succ R) = P(S \succ R \mid R \succ .4S + .6R)$. This assumption will not hold if subjects have different preferences—variation in preferences will induce a correlation. However, the representative agent stochastic choice model used here requires it, because it assumes identical preferences for all subjects.
24. The weighting function in (4) is only monotonically increasing when $\gamma > .27$ (otherwise, *higher* probabilities are perversely given *lower* weight). In unrestricted estimation, we got estimates less than .27 in three studies—Bernasconi's (1994) hypotenuse pairs ($\gamma = .02$), Conlisk's (1987) ($\gamma = .18$ for separable PT), and Gigliotti and Sopher's (1993) treatments 1 and 3 ($\gamma = .24$ for cumulative PT). Since the unrestricted maximum likelihoods were very close to the maximum likelihoods with $\gamma > .27$, the sensible restriction $\gamma > .27$ did not reduce likelihood much.
25. If EU is true, then -2 times the difference in log likelihoods between each theory and EU has approximately a chi-squared distribution with 1 degree of freedom.
26. Through the free parameter β, DAT effectively raises and lowers probability weights (as shown in equation (1) above), making $u(B)$ greater or less than $pu(S) + (1-p)u(R)$. In the single-agent approach with stochastic choice, a theory needs to account for the large percentage of subjects who obey betweenness *and* the smaller percentage of violations. If violations are not too common but are mostly quasi-concave, say, then the data are fit well by a theory like DAT, which makes $u(B)$ slightly larger than $pu(S) + (1-p)u(R)$.

However, DAT can never make $u(B)$ higher than *both* $u(S)$ and $u(R)$ (since it obeys betweenness), so it can never explain data in which a majority of people exhibit quasi-concavity. Indeed, in studies where more than half the subjects violated betweenness in the same direction (Conlisk, 1987; Prelec, 1990), DAT fit substantially worse than PT and CPT, and only a bit better than EU.

27. An example might illustrate how the two differ in our data, and why separable weighting sometimes works better.

Consider a triple of gambles from Bernasconi (1994, set 3 alternative A): $S = (.8, £12; .2, 0)$, $R = (.6, £20; .4, 0)$, and $B = .05R + .95S = (.03, £20; .76, £12; .21, 0)$. Separable and rank-dependent weights both give $u(S) = w(.8)\, u\,(£12)$ and $u(R) = w(.6)\, u\,(£20)$ (assuming $u(0) = 0$ for simplicity). Separable weights give

$$u(B) = w(.03)u(£20) + w(.76)u(£12) \tag{7}$$

and rank-dependent weights give

$$u(\mathrm{B}) = w(.03)u(£20) + [w(.79) - w(.03)]u(£12) \tag{8}$$

Triples in which B mixes two edge gambles, giving lower probability to S, usually exhibit a large fraction of quasi-concave patterns. To explain them a theory must make $u(B)$ large. Overweighting the low probability .03 raises $u(B)$, but does so equally in both separable and rank-dependent weighting (compare (7) and (8)). Overweighting .03 also lowers the incremental weight on u (£12) in the rank-dependent approach, $[w(.79) - w(.03)]$, but has no effect on the separable weight $w(.76)$. If the weighting function $w(\bullet)$ is steeper between 0 and .03 than between .76 and .79 (as in figure 7) then $w(.76)$ will be greater than $w(.79) - w(.03)$. The separable approach will then weight $u(£12)$ more heavily, compute a higher value for $u(B)$, and explain observed quasi-concavity more easily than the rank-dependent approach. (For example, for $\gamma = .55$, $w(.79) - w(.03) = .43$ and $w(.76) = .52$.)

References

Allais, Maurice. (1953). "Le Comportement de L'homme Rationel Devant le Risque, Critique des Postulates et Axiomes de L'ecole Americaine," *Econometrica* 21, 503–546.

Battalio, Ray C., John H. Kagel, and Jiranyakul Komain. (1990). "Testing Between Alternative Models of Choice Under Uncertainty: Some Initial Results," *Journal of Risk and Uncertainty* 3, 25–50.

Becker, Gordon, Morris DeGroot, and Jacob Marschak. (1963). "An Experimental Study of Some Stochastic Models for Wagers," *Behavioral Science* 3, 199–202.

Bernasconi, Michele. (1994). "Nonlinear Preference and Two-Stage Lotteries: Theories and Evidence," *The Economic Journal* 104, 54–70.

Bordley, Robert. (1992). "An Intransitive Expectations-Based Bayesian Variant of Prospect Theory," *Journal of Risk and Uncertainty* 5, 127–144.

Bordley, Robert, and Gordon Hazen. (1991). "SSB and Weighted Linear Utility as Expected Utility with Suspicion," *Management Science* 37, 396–408.

Camerer, Colin F. (1989). "An Experimental Test of Several Generalized Utility Theories," *Journal of Risk and Uncertainty* 2, 61–104.

Camerer, Colin F. (1992). "Recent Tests of Generalized Utility Theories." In W. Edwards (ed.), *Utility: Measurement, Theories, and Applications.* Dordrecht: Kluwer Academic, pp. 207-251.

Chew, Soo-Hong, and Kenneth R. MacCrimmon. (1979). "Alpha-nu Choice Theory: An Axiomatization of Expected Utility," Faculty of Commerce Working Paper No. 669, University of British Columbia.

Chew, Soo-Hong. (1983). "A Generalization of the Quasilinear Mean With Applications to the Measurement of the Income Inequality and Decision Theory Resolving the Allais Paradox," *Econometrica* 53, 1065–1092.

Chew, Soo-Hong, and William S. Waller. (1986). "Empirical Tests of Weighted Utility Theory," *Journal of Mathematical Psychology* 30, 55–72.

Chew, Soo-Hong. (1989). "Axiomatic Utility Theories With the Betweenness Property," *Annals of Operations Research* 19, 273–298.

Chew, Soo-Hong, and Larry G. Epstein. (1990). "A Unifying Approach to Axiomatic Non-Expected Utility Theories," *Journal of Economic Theory* 49, 207-240.

Chew, Soo-Hong, Larry G. Epstein, and Uzi Segal. (1991). "Mixture Symmetry and Quadratic Utility," *Econometrica* 59, 139-163.

Chew, Soo-Hong, Edi Karni, and Zvi Safra. (1987). "Risk Aversion in the Theory of Expected Utility with Rank Dependent Probabilities," *Journal of Economic Theory* 42, 370-381.

Conlisk, John. (1987). "Verifying the Betweenness Axiom With Questionnaire Evidence, or Not: Take Your Pick," *Economics Letter,* 25, 319-322.

Coombs, Clyde, and Lily Huang. (1976). "Tests of the Betweenness Property of Expected Utility," *Journal of Mathematical Psychology* 13, 323-337.

Cox, James, Vernon L. Smith, and James M. Walker. (1983). "A Test That Discriminates Between Two Models of the Dutch-first Auction Non-isomorphism," *Journal of Economic Behavior and Organization* 4, 205-219.

Crawford, Vincent P. (1988). "Stochastic Choice with Quasiconcave Preference Functions." University of California-San Diego Department of Economics Working Paper.

Crawford, Vincent P. (1990). "Equilibrium Without Independence," *Journal of Economic Theory* 50, 127-154.

Debreu, Gerard. (1952). "A Social Equilibrium Existence Theorem," *Proceedings of the National Academy of Sciences* 38, 886-893. Reprinted in *Mathematical Economics: Twenty Papers of Gerard Debreu.* (1983). New York: Cambridge University Press.

Dekel, Eddie. (1986). "An Axiomatic Characterization of Preference Under Uncertainty: Weakening the Independence Axiom," *Journal of Economic Theory* 40, 304-318.

Epstein, Larry G., and Stanley E. Zin. (1989). "Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica* 57, 937-969.

Epstein, Larry G., and Stanley E. Zin. (1991a). "Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis," *Journal of Political Economy* 99, 263-286.

Epstein, Larry G., and Stanley E. Zin. (1991b). "The Independence Axiom and Asset Returns," Department of Economics, University of Toronto.

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

Fishburn, Peter. (1983). "Transitive Measurable Utility," *Journal of Economic Theory* 31, 293-317.

Fishburn, Peter. (1988). *Nonlinear Preference and Utility Theory.* Baltimore: Johns Hopkins University Press.

Gigliotti, Gary, and Barry Sopher. (1993). "A Test of Generalized Expected Utility Theory," *Theory and Decision* 35, 75-106.

Green, Jerry, and Bruno Jullien. (1988). "Ordinal Independence in Nonlinear Utility Theory," *Journal of Risk and Uncertainty* 1, 355-387. (erratum in 2(1988), 119).

Gul, Faruk, and Otto Lantto. (1990). "Betweenness Satisfying Preferences and Dynamic Choice," *Journal of Economic Theory* 52, 162-177.

Gul, Faruk. (1991). "A Theory of Disappointment Aversion," *Econometrica,* 59 667-686.

Handa, Jagdish. (1977). "Risk, Probabilities, and a New Theory of Cardinal Utility," *Journal of Political Economy* 85, 97-122.

Harless, David. (1992). "Experimental Tests of Quasi-Bayesian Generalizations of Expected Utility Theory," Virginia Commonwealth University Department of Economics Working Paper.

Harless, David, and Colin Camerer. (in press). "The Predictive Utility of Generalized Utility Theories," *Econometrica.*

Hey, John and Chris Orme. (in press). "Investigating Parsimonious Generalizations of Expected Utility Theory Using Experimental Data," *Econometrica.*

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

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

Karni, Edi, and Zvi Safra. (1989a). "Ascending Bid Auctions with Behaviorally Consistent Bidders," *Annals of Operations Research* 19, 435-446.

Karni, Edi, and Zvi Safra. (1989b). "Dynamic Consistency, Revelations in Auctions and the Structure of Preferences," *Review of Economic Studies* 56, 421-434.

Kreps, David M., and Evan L. Porteus. (1978). "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica* 46, 185-200.

Kreps, David M., and Evan L. Porteus. (1979). "Temporal von Neumann-Morgenstern and Induced Preferences," *Journal of Economic Theory* 20, 81–109.

Lattimore, P.M., J.R. Baker, and A. Dryden Witte. (1992). "The Influence of Probability on Risky Choice," *Journal of Economic Behavior and Organization* 17, 377–400.

Loomes, Graham, and Robert Sugden. (1987). "Some Implications of a More General Form of Regret Theory," *Journal of Economic Theory* 41, 270–287.

Luce, R. Duncan. (1990). "Rational Versus Plausible Accounting Equivalences in Preference Judgments," *Psychological Science* 1, 225–234.

Luce, R. Duncan. (1991). "Rank- and Sign-Dependent Linear Utility Models for Binary Gambles," *Journal of Economic Theory* 53, 75–100.

Luce, R. Duncan, and Peter 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. Duncan, and Patrick Suppes. (1965). "Preference, Utility, and Subjective Probability." In R.D. Luce, R.B. Bush, and E. Galanter (eds.), *Handbook of Mathematical Psychology*, Vol. 3. New York: Wiley, pp. 249–410.

Machina, Mark. (1982). " 'Expected Utility' Analysis without the Independence Axiom," *Econometrica* 50, 277–323.

Machina, Mark. (1985). "Stochastic Choice Functions Generated from Deterministic Preferences Over Lotteries," *Economic Journal* 95, 575–594.

Machina, Mark. (1987). "Choice Under Uncertainty: Problems Solved and Unsolved," *Journal of Economic Perspectives* 1, 121–154.

Machina, Mark. (1989). "Dynamic Consistency and Non-Expected Utility Models of Choice Under Uncertainty," *Journal of Economic Literature* 26, 1622–1668.

Marschak, Jacob. (1950). "Rational Behavior, Uncertain Prospects, and Measurable Utility," *Econometrica* 18, 111–141.

Prelec, Drazěn. (1990). "A 'Pseudo-endowment' Effect, and its Implications for Some Recent Nonexpected Utility Models," *Journal of Risk and Uncertainty* 3, 247–259.

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

Röell, Alissa. (1987). "Risk Aversion in Quiggin and Yaari's Rank-Order Model of Choice under Uncertainty," *Economic Journal* 97, 143–159.

Schmeidler, David. (1989). "Subjective Expected Utility without Additivity," *Econometrica*, 57 571–587.

Segal, Uzi. (1989). "Anticipated Utility: A Measure Representation Approach," *Annals of Operations Research* 19, 359–373.

Segal, Uzi. (1990). "Two-stage Lotteries Without the Reduction Axiom," *Econometrica* 58, 349–377.

Tversky, Amos, and Daniel Kahneman. (1986). "Rational Choice and the Framing of Decisions," *Journal of Business* 59, S251–S278. Reprinted in R. Hogarth and M. Reder (eds.), *Rational Choice: The Contrast between Economics and Psychology*. (1987). Chicago: University of Chicago Press.

Tversky, Amos, Paul Slovic, and Daniel Kahneman. (1990). "The Causes of Preference Reversal," *The American Economic Review* 80, 204–217.

Tversky, Amos, and Daniel Kahneman. (1992). "Advances in Prospect Theory: Cumulative Representations of Uncertainty," *Journal of Risk and Uncertainty* 5, 297–323.

Viscusi, W. Kip. (1989). "Prospective Reference Theory: Toward An Explanation of the Paradoxes," *Journal of Risk and Uncertainty* 2, 235–264.

Weber, Robert J. (1982). "The Allais Paradox, Dutch Auctions, and Alpha-utility Theory," MEDS Department Discussion Paper No. 536, Northwestern University.

Weber, Martin, and Colin Camerer. (1987). "Recent Developments in Modelling Preferences Under Risk," *OR Spektrum* 9, 129–151.

Wu, George. (1993). "Editing and Prospect Theory: Ordinal Independence and Outcome Cancellation," Harvard Business School Managerial Economics Working Paper.

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