Economics 142: Choice under Uncertainty (or Certainty) **Winter 2008**
Vincent Crawford (with very large debts to Matthew Rabin and especially Botond Koszegi)

Background: Classical theory of choice under certainty

Rational choice (complete, transitive, and continuous preferences) over certain outcomes and representation of preferences via maximization of an ordinal utility function of outcomes.

The individual makes choices “as if” to maximize the utility function; utility maximization is just a compact, tractable way for us to describe the individual’s choices in various settings.

We can view the utility function as a compact way of storing intuition about behavior from simple experiments or though-experiments and transporting it to new situations.

The preferences represented can be anything—self-interested or not, increasing in intuitive directions (more income or consumption) or not—although there are strong conventions in mainstream economics about what they are normally defined over—own income or consumption rather than both own and others’, levels of final outcomes rather than changes.

Thus if you think the mainstream approach is narrow or wrong-headed, it may make as much or more sense to complain about those conventions than about the idea of rationality per se.

Background: Classical "expected utility" theory of choice under uncertainty

This is the standard way to describe people's preferences over uncertain outcomes. The Marschak reading on the reading list, linked on the course page, is a readable introduction.

The basic idea is that if an individual's preferences satisfy certain axioms, discussed below, and the uncertainty is over which of a given list of outcomes will happen, then a person's preferences over probability distributions over those outcomes can be described (much as for certain outcomes, although there is an important difference) by assigning utility numbers (called "von Neumann-Morgenstern utilities" in analyses of individual decisions or, equivalently, "payoffs" in games), one to each possible outcome, and assuming that the person chooses as if to maximize "expected utility"—the mathematical expectation of the utility of the realized outcome.

Example: Suppose that your initial lifetime wealth w is \$2 million dollars, you are asked to choose whether or not to accept a bet (investment opportunity, insurance contract, etc.) that will add either x, y, or z (which could be negative) to your wealth, with respective probabilities p, q, or $1 - p - q$. Suppose further that you care only about your final lifetime wealth.

Then the claim is that, under the axioms mentioned about, the analyst can assign utilities to the possible final outcomes w, $w + x$, $w + y$, and $w + z$, call them $u(w)$, $u(w + x)$, $u(w + y)$, and $u(w + z)$, such that the person will accept the bet if and only if (ignoring ties) $u(w) < pu(w + x) + qu(w + y) + (1 - p - q)u(w + z)$. (In other words, if the expected utility of "w for sure" is less than that of a random addition to w of x, y, or z with probabilities p, q, or $1 - p - q$.) The vN-M utility function whose expectation the individual acts as if to maximize is a compact way to describe the individual's choices in various settings involving uncertainty.

There are two important assumptions in this example:

● That the person's preferences are "well-behaved" enough to be represented by a "preference function" (so called to distinguish it from the utility function) over probability distributions that (with outcomes fixed and distributions over the fixed outcomes described by lists of probabilities) is linear in the probabilities (that is, it is an expected utility).

● That the person's preferences respond only to final, level (as opposed to change) outcomes, in this case of the person's own lifetime wealth.

The first assumption is logically justified by a famous result known as the Von Neumann-Morgenstern Theorem.

The theorem's axioms are sometimes systematically violated in observed behavior, and the axioms that the theorem uses to justify expected-utility maximization are not completely uncontroversial. But these violations seem behaviorally and economically less important than violations of the second assumption.

The second assumption is not at all logically necessary to use expected-utility maximization to describe choice under uncertainty.

It is only a convention of mainstream economics, which could be replaced by an alternative convention to yield an alternative expected-utility characterization of choice under uncertainty, as we shall do below.

First let's record the logic of the first assumption. (This snapshot and others in this section are from Machina, "Expected Utility Hypothesis," linked on the course web page.)

Recall that in such a case the objects of choice consist of all probability distributions $P = (p_1, \ldots, p_n)$ over $\{x_1, \ldots, x_n\}$, so that the following axioms refer to the individuals' weak preference relation $\succsim$ over this set, where $P^* \succsim P$ is read 'P^* is weakly preferred (i.e. preferred or indifferent) to P' (the associated strict preference relation $\succ$ and indifference relation $\sim$ are defined in the usual manner):

Completeness: For any two distributions P and P^* either $P^* \succsim P$, $P \succsim P^*$, or both.

Transitivity: If $P^{**} \succsim P^*$ and $P^* \succsim P$, then $P^{**} \succsim P$.

Mixture Continuity: If $P^{**} \succsim P^* \succsim P$, then there exists some $\lambda \in [0, 1]$ such that $P^* \sim \lambda P^{**} + (1-\lambda)P$, and

Independence: For any two distributions P and P^*, $P^* \succsim P$ if and only if $\lambda P^* + (1-\lambda)P^{**} \succsim \lambda P + (1-\lambda)P^{**}$ for all $\lambda \in (0, 1]$ and all P^{**},

where $\lambda P + (1-\lambda)P^*$ denotes the 'probability mixture' of P and P^*, i.e., the lottery with probabilities

$$(\lambda p_1 + (1-\lambda)p_1^*, \ldots, \lambda p_n + (1-\lambda)p_n^*).$$

Von Neumann-Morgenstern Theorem: Complete, transitive, and continuous preferences over probability distributions of outcomes that satisfy the "independence axiom" can be represented by the maximization of "expected utility" (just as complete, transitive, and continuous preferences over certain outcomes can be represented by the maximization of standard utility).

There are two main differences from the classical theory of choice under certainty:

(i) "Outcomes" are now probability distributions (over a prespecified set of certain outcomes, which are included as degenerate probability distributions).

(ii) The theory imposes a specific restriction on preferences ("Independence"). Independence implies that the preference function over probability distributions that the individual maximizes is linear in the probabilities (not the same as the vN-M utility function u(x) being linear in x!).

What does the Independence Axiom say? Consider a Spinner: Prob(Red) = 89%, Prob(Blue) = 10%, Prob(Yellow) = 1%. Suppose I asked you to choose between two lotteries:

L_A: x if Red, y if Blue, and z if Yellow.
L_B: x if Red, w if Blue orYellow.

The Independence Axiom says that you don't need to know x to choose between L_A and L_B: Independence is like separability of preferences across states (Red versus Blue or Yellow).

The vN-M Theorem also assumes the probability distributions are objective or at least known.

Leonard Savage (*The Foundations of Statistics*) generalized the theory to allow subjective probabilities, showing (roughly speaking) that not "knowing" the probabilities matters only when you can take actions (testing, search, etc.) to learn about them.

Linearity of expected-utility indifference curves in (p_1, p_3) space:

The formal representation of the objects of choice, and hence of the expected utility preference function, depends upon the structure of the set of possible outcomes. When there are a finite number of outcomes $\{x_1, \ldots, x_n\}$, we can represent any probability distribution over this set by its vector of probabilities $P = (p_1, \ldots, p_n)$ (where $p_i = \text{prob}(\tilde{x} = x_i)$), and the preference function takes the form

$$V(P) = V(p_1, \ldots, p_n) \equiv \Sigma U_i p_i.$$

lotteries. Since every probability distribution (p_1, p_2, p_3) over this set must satisfy the condition $\Sigma p_i = 1$, we may represent each such distribution by a point in the unit triangle in the (p_1, p_3) plane, with p_2 given by $p_2 = 1 - p_1 - p_3$ (Figures 1 and 2). Since they represent the loci of solutions to the equations

$$U_1 p_1 + U_2 p_2 + U_3 p_3 = U_2 - [U_2 - U_1] \cdot p_1 + [U_3 - U_2] \cdot p_3$$
$$= \text{constant}$$

for the fixed utility indices $\{U_1, U_2, U_3\}$, the indifference curves of an expected utility maximizer consist of parallel straight lines in the triangle of slope $[U_2 - U_1]/[U_3 - U_2]$, as illustrated by the solid lines in Figure 1. An example of indifference curves which

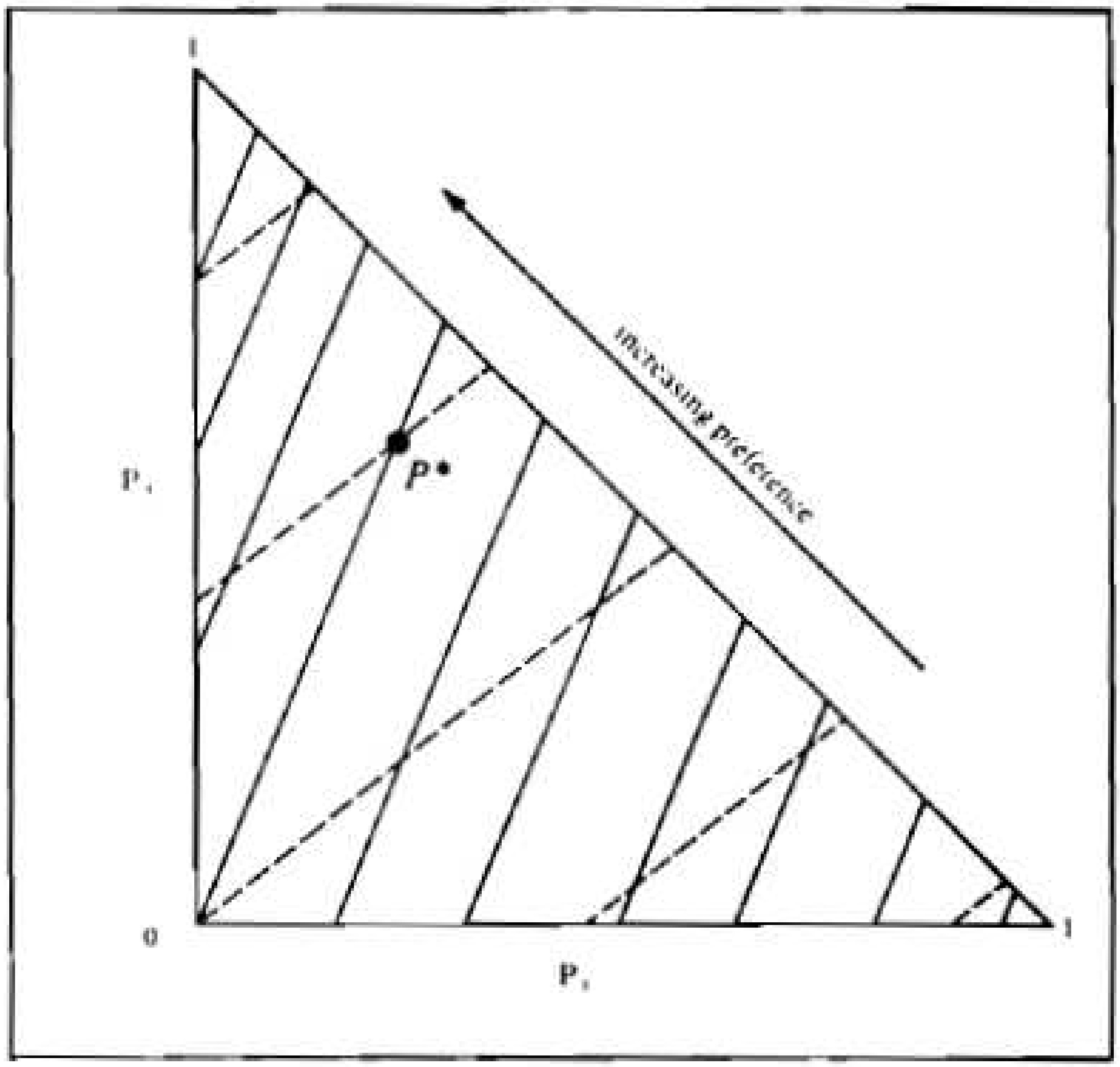


Figure 1 Expected Utility Indifference Curves

Possible violations of the von Neumann-Morgenstern axioms

As mentioned above, observed behavior sometimes systematically violates the von Neumann-Morgenstern Theorem's axioms. There are two main kinds of violation:

- Violations of Independence/Nonlinearity of the preference function in the probabilities
- Ambiguity aversion/Dependence of preference over probability distributions on source

Although these violations seem behaviorally and economically less important than violations of the mainstream convention that the person's preferences respond only to final, level (as opposed to change) outcomes, it is important to know about them.

Violations of Independence/Nonlinearity of the preference function in the probabilities

Suppose you are separately offered the following two choices:

- a_1 vs. a_2
- a_3 vs. a_4

where the a_i are defined as follows:

	a_1	a_2	a_3	a_4
Prob {\$0}	0.00	0.01	0.90	0.89
Prob {\$1 million}	1.00	0.89	0.00	0.11
Prob {\$5 million}	0.00	0.10	0.10	0.00

Record your choices, one each (independently) for a_1 vs. a_2 and for a_3 vs. a_4.

Allais Paradox: Most people choose a_3 over a_4, but a_1 over a_2. This seems reasonable (to me), but it violates the Independence Axiom because moving (in probability space) from a_1 to a_2 is the same (in direction and magnitude) as moving from a_4 to a_3 (note order!). So with preferences that are linear in the probabilities the moves should either be both good or both bad.

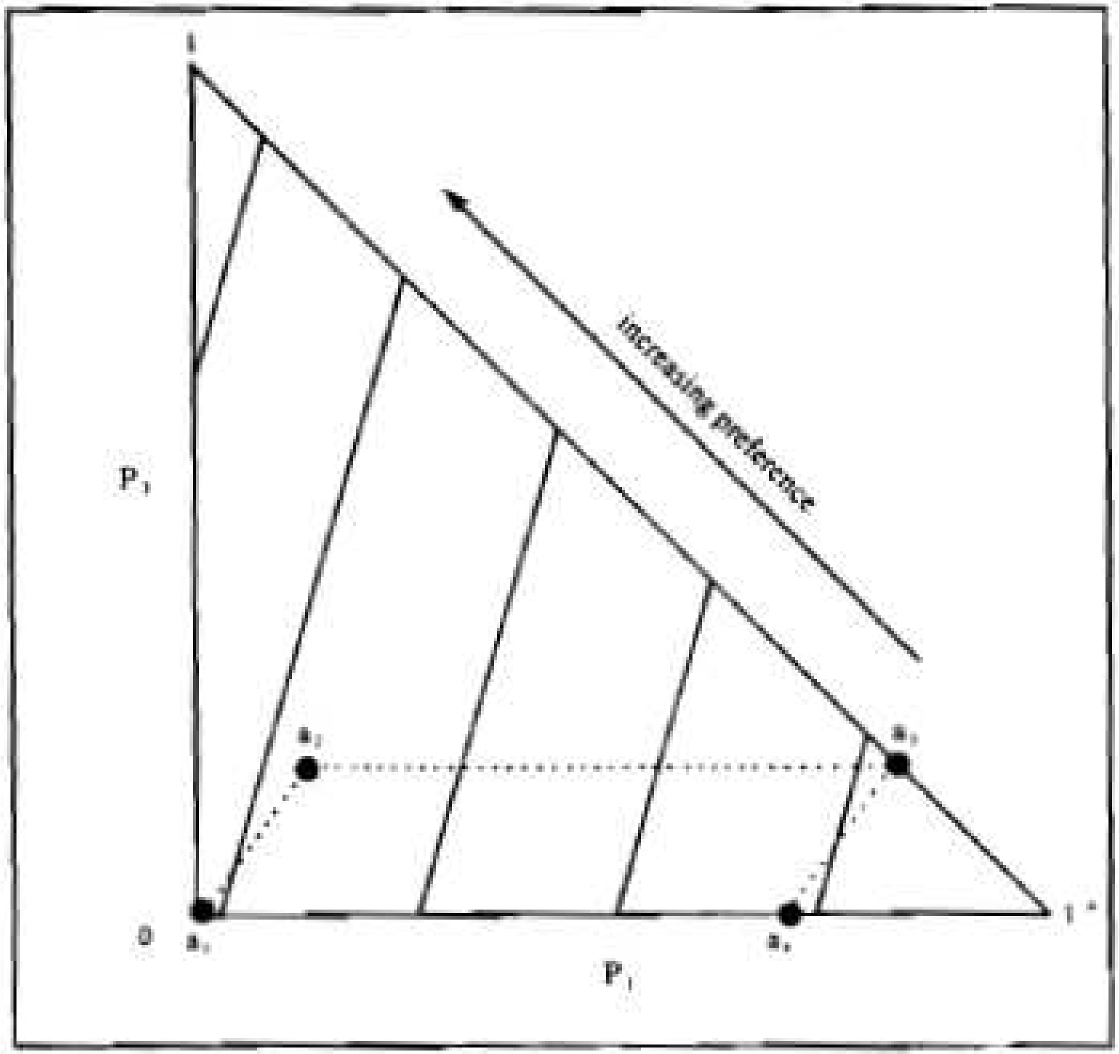


Figure 5 Allais Paradox with Expected Utility Indifference Curves

Ambiguity aversion/Dependence of preference over probability distributions on source

The Allais Paradox raises doubts about Independence/Linearity in the probabilities, but not necessarily about whether preferences over uncertain outcomes are independent of the kind or source of the uncertainty.

The Ellsberg Paradox suggests that either people don't form coherent subjective probabilistic assessments, or the way probabilities enter their preferences depends on the kind or source of the uncertainty from which those probabilities are derived:

Suppose you are told (and you believe) that Urn 1 contains 100 balls. Each of them is certainly either Red or Black, but you are not told how many Red and Black balls there are.

However, you are told (and you believe) that Urn 2 contains exactly 50 Red and 50 Black balls.

You are separately offered the following two choices:

- You must choose one of two gambles, either
  - a_1: Pick a ball from Urn 1, win \$100 if it's Red, \$0 if it's Black.
  - a_2: Pick a ball from Urn 2, win \$100 if it's Red, \$0 if it's Black.

- You must choose one of two gambles, either
  - b_1: Pick a ball from Urn 1, win \$100 if it's Black, \$0 if it's Red.
  - b_2: Pick a ball from Urn 2, win \$100 if it's Black, \$0 if it's Red.

Ellsberg Paradox: Most people choose a_2 over a_1, but choose b_2 over b_1.

Again this seems reasonable: the choices of a_2 over a_1 and b_2 over b_1 could both reflect aversion to uncertainty about the probabilities ("ambiguity aversion").

But (assuming that such choices aren't just accidental consequences of widespread indifference, and that people prefer more money to less), choosing a_2 over a_1 but b_2 over b_1 cannot be consistent with maximizing a preference function that depends only on the probabilities.

For, if preferences depend only on the probabilities, the majority choice of a_2 over a_1 "reveals" that the chooser thinks the expected number of Red balls in Urn 1 (for her/his subjective probability distribution, whatever it is) is less than 50, the known number of Red balls in Urn 2.

But the majority choice of b_2 over b_1 also "reveals" that the chooser thinks the expected number of Red balls in Urn 1 is more than 50, a contradiction.

(This conclusion does not require Independence: Anyone who likes money, even with nonlinear preference function, wants to maximize the probability of winning $100 rather than $0.)

The contradiction shows that preferences cannot depend only on the probabilities.

Conventional assumptions about preferences, “liking money,” and risk aversion

To give the theory empirical content, we need to add assumptions on what preferences are about.

Just as with the theory of choice under certainty, unless we are willing to commit to a particular specification of what the individual cares about, and to some extent how, the theory is flexible enough to allow (almost) anything, and is therefore (almost) useless.

Let’s assume (for illustration; not essential to the theory) that the individual cares only about his money income, independent of the “state” that describes how the uncertainty is resolved.

(E.g. in roulette, let’s assume the individual cares only about money winnings, not how he wins, e.g. on red versus black, or on some particular number. If he did care how he won, we could still use expected utility maximization to describe his choices by including how he won in the description of an outcome (distinguishing the outcomes “win $10 betting on red” and “win $10 betting on black.” But then the theory would make much weaker predictions because there would be no necessary relation between his attitudes toward risk betting on red and betting on black.)

Let’s also assume that the person prefers more money to less, and (so we can graph things) that there are only three possible money outcomes.

Preference for first-order stochastic dominance (generalization of "liking money" to uncertain money outcomes) and upward-sloping indifference curves in (p_1, p_3) space:

When the outcomes $\{x_1, x_2, x_3\}$ represent different levels of wealth with $x_1 < x_2 < x_3$, this diagram can be used to illustrate other possible aspects of an expected utility maximizer's attitudes toward risk. On the general principle that more wealth is better, it is typically postulated that any change in a distribution (p_1, p_2, p_3) which increases p_3 at the expense of p_2, increases p_2 at the expense of p_1, or both, will be preferred by the individual: this property is known as 'first-order stochastic dominance preference'. Since such shifts of probability mass are represented by north, west or north-west movements in the diagram, first-order stochastic dominance preference is equivalent to the condition that indifference curves are upward sloping, with more preferred indifference curves lying to the north-west. Algebraically, this is equivalent to the condition $U_1 < U_2 < U_3$.

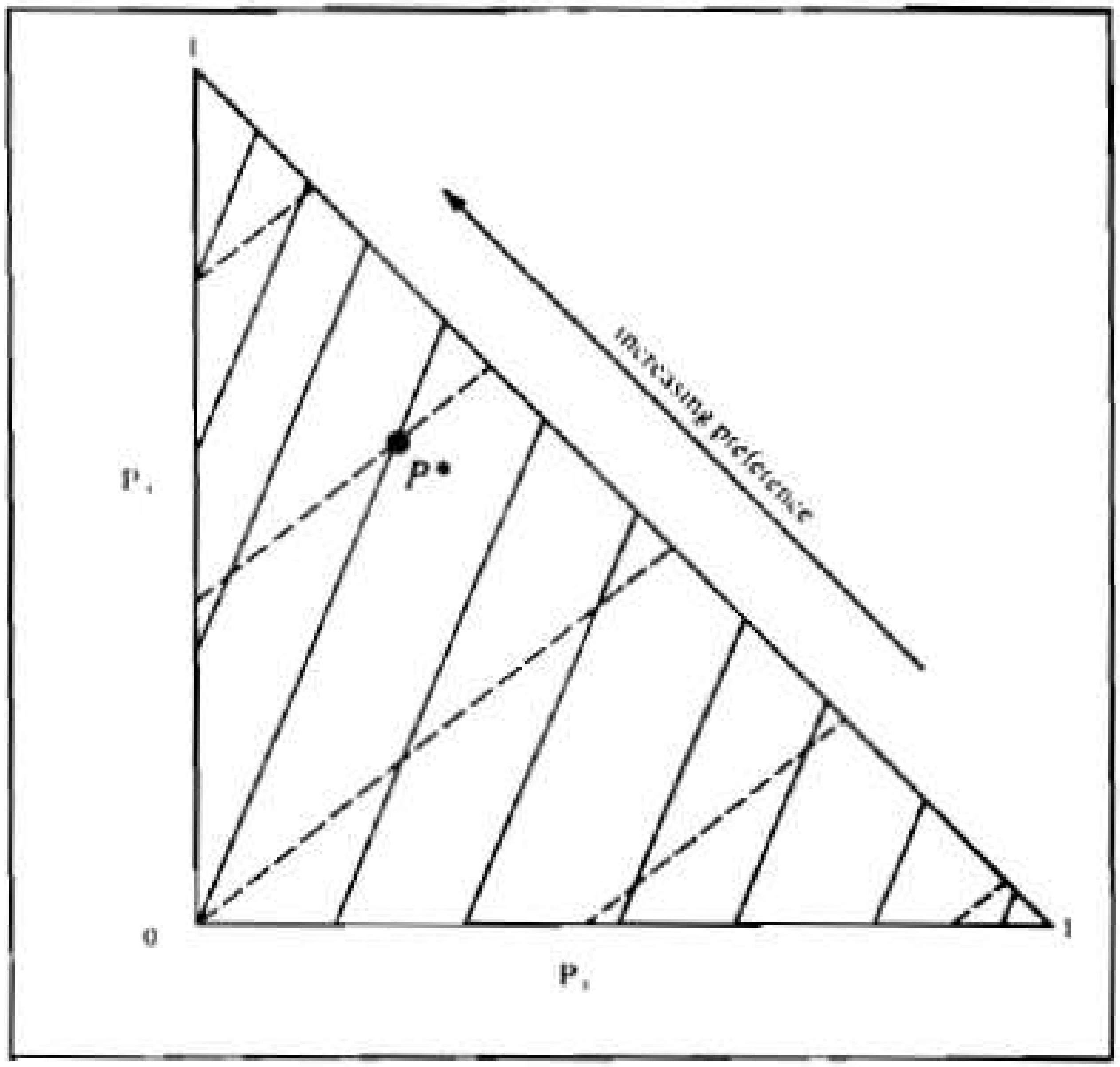


Figure 1 Expected Utility Indifference Curves

Risk aversion and slopes of indifference curves in (p_1, p_3) space

Another widely (though not universally hypothesized aspect of attitudes towards risk is that of 'risk aversion' (e.g. Arrow, 1974, ch. 3; Pratt, 1964). To illustrate this property of preferences, consider the dashed lines in Figure 1, which represent loci of solutions to the equations

$$x_1 p_1 + x_2 p_2 + x_3 p_3 = x_2 - [x_2 - x_1] \cdot p_1 + [x_3 - x_2] \cdot p_3$$
$$= \text{constant}$$

and hence may be termed 'iso-expected *value* loci'. Since north-east movements along any of these loci consist of increasing the tail probabilities p_1 and p_3 at the expense of middle probability p_2 in a manner which preserves the mean of the distribution, they correspond to what are termed 'mean preserving increases in risk' (e.g. Rothschild and Stiglitz, 1970, 1971). An individual is said to be 'risk averse' if such increases in risk always lead to less preferred indifference curves, which is equivalent to the graphical condition that the indifference curves be steeper than the iso-expected value loci. Since the slope of the latter is given by $[x_2 - x_1]/[x_3 - x_2]$, this is equivalent to the algebraic condition that $[U_2 - U_1]/[x_2 - x_1] > [U_3 - U_2]/[x_3 - x_2]$. Conversely, individuals who *prefer* mean preserving increases in risk are termed 'risk loving': such individuals' indifference curves will be flatter than the iso-expected value loci, and their utility indices will satisfy $[U_2 - U_1]/[x_2 - x_1] < [U_3 - U_2]/[x_3 - x_2]$.

Note that the ideas of risk aversion and FOSD are more general than linear indifference curves.

Risk aversion and concavity of the von Neumann-Morgenstern utility function

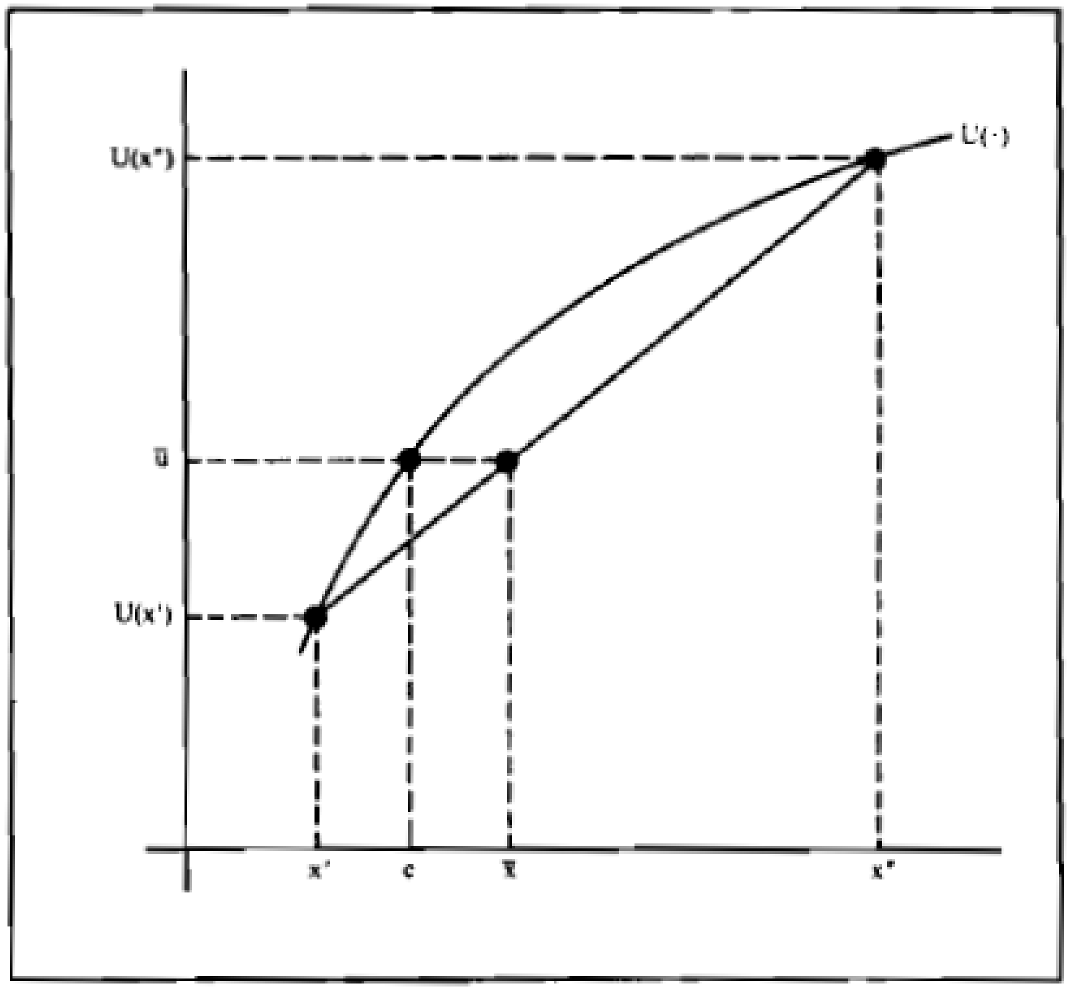


Figure 3 Von Neumann-Morgenstern Utility Function of a Risk Averse Individual

Evidence for reference-dependent preferences

Now let's consider the evidence on the second, conventional assumption that the person's preferences respond only to final, level (as opposed to change) outcomes.

(See the Kahneman readings in the first section of the reading list for psychological background.)

First re consider the answers to questions 1 (a and b) of the survey.

1a. Would you choose to lose $500 for sure or to lose $1000 with probability 0.5?

1b. Would you choose to receive $500 for sure or to receive $1000 with probability 0.5?

● Most people who answer questions like 1a choose to lose $1000 with probability 0.5 rather than losing $500 for sure, suggesting "risk-loving" behavior with respect to losses. (This suggests that people dislike losses so much they are willing to take a fairly large, equal-expected-money-outcome risk just to reduce the probability of a loss.)

● By contrast, most people who answer questions like 1b chose to receive $500 for sure rather than $1000 with probability 0.5, suggesting "risk-averse" behavior with respect to gains.

But if preferences are defined over absolute, final levels of outcomes rather than changes, choices should be qualitatively the same—both risk-loving or risk-averting—for 1a and 1b.

In the Introduction I suggested—and then dismissed—the possibility that such a large flip could plausibly be explained by income effects, even though those who answered 1b are on average richer (because of the gains) than those who answered 1a.

Income effects can be tested as an alternative explanation by a different version of the experiment done (in this case with hypothetical payoffs) by Kahneman and Tversky:

Problem number 1: In addition to whatever you own, you have been given $1000.

You are now asked to choose between A: receiving another $1000 with probability 0.5 and B: receiving another $500 for sure. (84% chose B.)

Problem number 2: In addition to whatever you own, you have been given $2000.

You are now asked to choose between C: losing $1000 with probability 0.5 and D: losing $500 for sure. (69% chose C.)

In terms of the probability distributions of final outcomes, these two choices are mathematically identical.

Thus the large flip in the distribution of choices must be somehow due to the change in perspective. A plausible hypothesis is that problem 1's framing makes people think of it as a choice between gains, while problem 2's makes people think of it as a choice between losses.

This appears to make people risk-averse in problem 1 but risk-loving in problem 2.

Mugs

In a famous experiment, Kahneman, Knetsch, and Thaler (1990, 1991) randomly gave mugs to half the subjects in a classroom experiment ("owners") and nothing to the others ("non-owners").

They then elicited selling prices for owners and buying prices for non-owners.

Here they used a procedure that gives subjects an incentive to reveal their true prices: Subjects are told that a price has been selected randomly, and is sealed in an envelope in the front of the room (in plain view of everyone). They then get a sheet of paper with a bunch of possible prices listed, and they are asked to indicate whether they would buy at each price. The highest price at which a buyer expresses a willingness to buy is taken as her/his "buying price," and the highest price at which a seller expresses a willingness to keep the good is taken as her/his "selling price."

If they had elicited prices from mug owners in the field, there might have been selection effects, in that we might expect mug owners to have higher prices than non-owners, on average, just because they were the ones who chose to acquire them.

We might also be concerned that owners knew more about mug quality than non-owners, etc.

But in the experiment, owners and non-owners were randomly assigned, and all had equal opportunity to inspect the mugs. Thus in a large enough sample, with a common value distribution, supply and demand "should" be mirror images of each other. But…

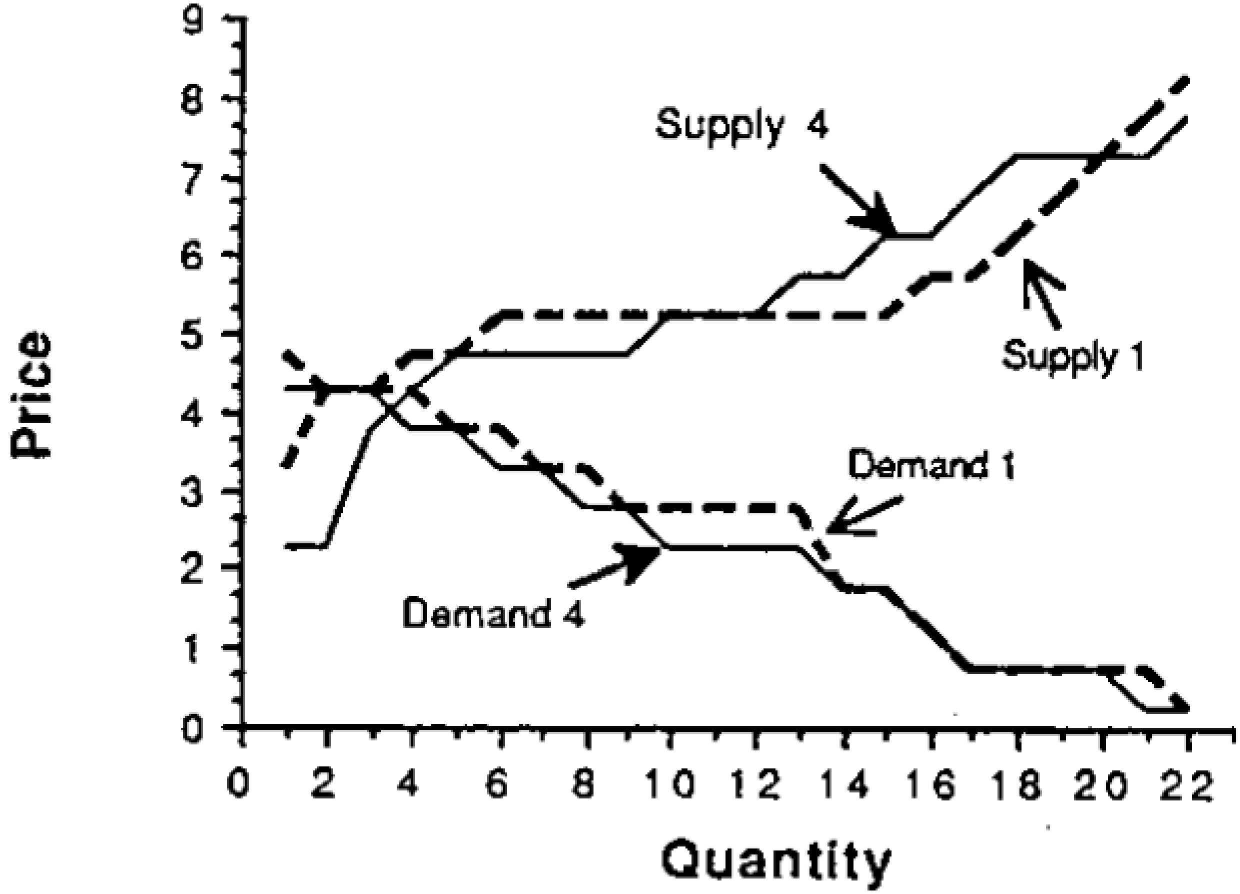


FIG. 1.—Supply and demand curves, markets 1 and 4

The average buying price of non-owners was about $3.50, and the average selling price of owners was about $7.00: way too big a gap to be random.

Again, this could conceivably be due to income effects, in that those subjects who received mugs were, on average, slightly richer than those who did not. But income effects from mugs are not likely to be large enough to explain a gap as big as that between $3.50 and $7.00.

Further, they can be ruled out by an experiment in which another group of subjects, “choosers,” are told they are going to be given either a mug or money, and that they should choose the amount of money that makes them indifferent between the mug and that amount of money.

Choosers have the same incentives that sellers in the original experiment did to reveal their true “reservation price” for the mug.

Yet in a typical experiment, the average selling, buying, and choosing prices were $7.12, $2.87, and $3.12 respectively. Thus choosers, who have approximately the same “income” as owner/sellers (because they know they are going to get either a mug or at least an equivalent amount of money), have reservation prices like buyer/nonowners, who have no such income.

The results of experiments like these seem to reflect reference-dependent preferences with “loss aversion.” Just as I suggested that in Survey questions 1 (a and b), gains and losses are defined relative to the status quo ante of $0, here we can imagine that mug owners treat having a mug as their reference point, and consider not having a mug to be a loss; and that non-owners treat not having a mug as their reference point, and consider getting a mug as a gain. If (as the evidence also suggests) people are more sensitive to losses than they are to same-size gains, then owner/sellers will have higher reservation prices than either nonowner/buyers or choosers.

vN-M expected utility maximization and aversion to small risks

More subtle evidence for reference-dependence is implicit in the widespread observation that people seem to be much more averse to small risks than the vN-M theory would predict, given their willingness to take larger risks. Rabin and Thaler (*JEP* 2001) make this point vividly:

place, however, we will show that this explanation for risk aversion is not plausible in most cases where economists invoke it.

To help see why we make such a claim, suppose we know that Johnny is a risk-averse expected utility maximizer, and that he will always turn down the 50-50 gamble of losing $10 or gaining $11. What else can we say about Johnny? Specifically, can we say anything about bets Johnny will be willing to accept in which there is a 50 percent chance of losing $100 and a 50 percent chance of winning some amount $Y? Consider the following multiple-choice quiz:

From the description above, what is the biggest Y such that we *know* Johnny will turn down a 50-50 lose $100/win $Y bet?

a) $110
b) $221
c) $2,000
d) $20,242
e) $1.1 million
f) $2.5 billion
g) Johnny will reject the bet no matter what Y is.
h) We can't say without more information about Johnny's utility function.

Before you choose an answer, we remind you that we are asking what is the highest value of Y making this statement true for *all* possible preferences consistent with Johnny being a risk-averse expected utility maximizer who turns down the 50/50 lose $10/gain $11 for all initial wealth levels. Make no ancillary assumptions, for instance, about the functional form of Johnny's utility function beyond the fact that it is an increasing and concave function of wealth. Stop now, and make a guess.

Did you guess a, b, or c? If so, you are wrong. Guess again. Did you guess *d*? Maybe you figured we wouldn't be asking if the answer weren't shocking, so you made a ridiculous guess like *e*, or maybe even *f*. If so, again you are wrong. Perhaps you guessed *h*, thinking that the question is impossible to answer with so little to go on. Wrong again.

The correct answer is *g*. Johnny will turn down *any* bet with a 50 percent risk of losing at least $100, no matter how high the upside risk.

Johnny would, of course, have to be insane to turn down bets like *d*, *e*, and *f*. So, what is going on here? In conventional expected utility theory, risk aversion comes *solely* from the concavity of a person's utility defined over wealth levels. Johnny's risk aversion over the small bet means, therefore, that his marginal utility for wealth must diminish incredibly rapidly. This means, in turn, that even the chance for staggering gains in wealth provide him with so little marginal utility that he would be unwilling to risk anything significant to get these gains.

Suppose Johnny is an expected-utility-of-wealth maximizer who would turn down a 50/50 lose \$1,000/gain \$1,100 bet (or similar risks) for a non-trivial range of initial wealth levels.

Claim: Empirically, the vast majority of people would turn down such bets if they were offered in isolation, and would do so over a huge range of given lifetime wealth levels.

Rejection of the \$1,000/\$1,100 bet based on diminishing marginal utility of wealth implies an over 9% drop in marginal utility of wealth with a \$2,100 increase in lifetime wealth. But this implies that marginal utility of wealth plummets for larger changes unless there are dramatic shifts in risk attitudes over larger changes in wealth.

Hence, in the absence of such dramatic shifts, turning down this bet means that Johnny's marginal utility for money would be at most 34% of his current marginal utility of wealth if he were \$21,000 wealthier ... and if Johnny became \$105,000 wealthier in lifetime wealth—which is something less than \$5,000 in pre-tax income per year, say—then he would value income only at most $\approx$ 0.8% ($\approx (10/11)^{50}$) as much as he currently does.

Such a plummet in marginal utility of wealth means incredible risk aversion over larger stakes. If Johnny's marginal utility of wealth drops by 99% when he is \$105,000 wealthier, for instance, then—even if he were risk-neutral above his current wealth level but averse to \$1,000/\$1,100 bets below his current wealth level—Johnny would turn down a 50/50 lose \$210,000/gain \$10 million bet at his current wealth level. And if Johnny were risk neutral above his current wealth level but averse to 50/50 lose \$10/gain \$11 bets below his current wealth level, then he would turn down a 50/50 lose \$22,000/gain \$100 billion bet.

Aside: People also seem to be more comforted by compounding small risks than vN-M theory predicts (Rabin-Thaler (*JEP* 2001) again:)

Expected utility theory's presumption that attitudes towards moderate-scale and large-scale risks derive from the same utility-of-wealth function relates to a widely discussed implication of the theory: that people have approximately the same risk attitude towards an aggregation of independent, identical gambles as towards each of the independent gambles. This observation was introduced in a famous article by Paul Samuelson (1963), who reports that he once offered a colleague a bet in which he could flip a coin and either gain \$200 or lose \$100. The colleague declined the bet, but announced his willingness to accept 100 such bets together. Samuelson showed that this pair of choices was inconsistent with expected utility theory, which implies that if (for some range of wealth levels) a person turns down a particular gamble, then the person should also turn down an offer to play many of those gambles.

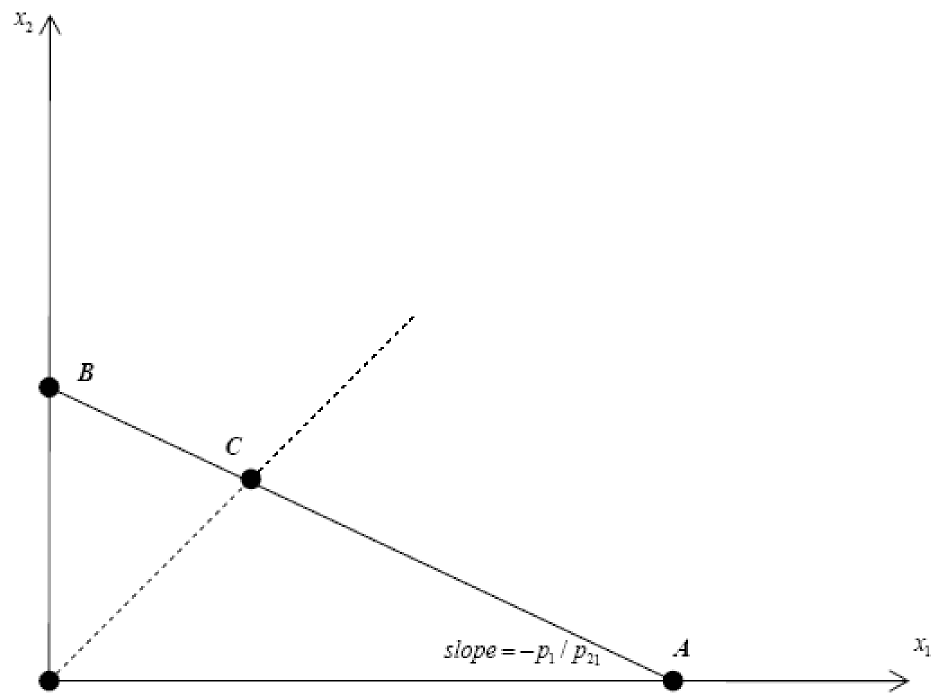
When Samuelson showed that his colleague's pair of choices was not consistent with expected utility theory, Samuelson thought that the mistake his colleague made was in accepting the aggregated bet, not in turning down the individual bet. This judgement is one we cannot share. The aggregated gamble of 100 50-50 lose \$100/gain \$200 bets has an expected return of \$5,000, with only a 1/2,300 chance of losing any money and merely a 1/62,000 chance of losing more than \$1,000. A good lawyer could have you declared legally insane for turning down this gamble.

By treating expected utility theory as a valid explanation of his colleague's aversion to the single gamble, and not questioning the plausibility of rejecting the aggregated gamble, we feel that Samuelson and economists since then have missed the true implications of his equivalence theorem. Samuelson and others have speculated as to the error his colleague was making, such as thinking that the variance of a repeated series of bets is lower than the variance of one bet (whereas, of course, the variance increases, though not proportionally, with repetition). Others have played off the fact that the equivalence theorem holds only approximately to explore the precise qualitative relationship that expected utility permits between risk attitudes over one draw and many independent draws of a bet. But our argument here reveals the irrelevance of these lines of reasoning. It does not matter what predictions expected utility theory makes about Samuelson's colleague, since the degree of risk aversion he exhibited proved he was not an expected utility maximizer. In fact, under exactly the same assumptions invoked by Samuelson, the theorem in Rabin (2000) implies that a risk-averse expected utility maximizer who turns down a 50-50 lose $100/gain $200 gamble will turn down a 50-50 lose $200/gain $20,000 gamble. This has an expected return of $9,900—with exactly zero chance of losing more than $200. Even a *lousy* lawyer could have you declared legally insane for turning down *this* gamble.

Aside: Implications for indifference maps in (s_1, s_2) space with state-independent, differentiable expected-utility preferences over money outcomes

There are two states, s_1 and s_2, and an individual who knows their probabilities, p_1 and p_2, chooses among state-contingent consumption bundles (x_1, x_2), facing budget as in Figure 1

Figure 1: An example of a budget constraint with two states and two assets.



(i) constant marginal rates of substitution equal to odds ratio on 45º line in (s_1, s_2) space:

On an indifference curve $p_1u(x_1) + p_2u(x_2)$ = constant.

Totally differentiating yields $p_1u'(x_1)\mathrm{d}x_1/\mathrm{d}x_2 + p_2u'(x_2) = 0$ or $\mathrm{d}x_1/\mathrm{d}x_2 = -p_2u(x_2)/p_1u'(x_1)$.

So on the 45º line $\mathrm{d}x_1/\mathrm{d}x_2 = -p_2/p_1$.

(ii) constant ratios of marginal rates of substitution across ends of rectangles oriented with the axes located anywhere in (s_1, s_2) space:

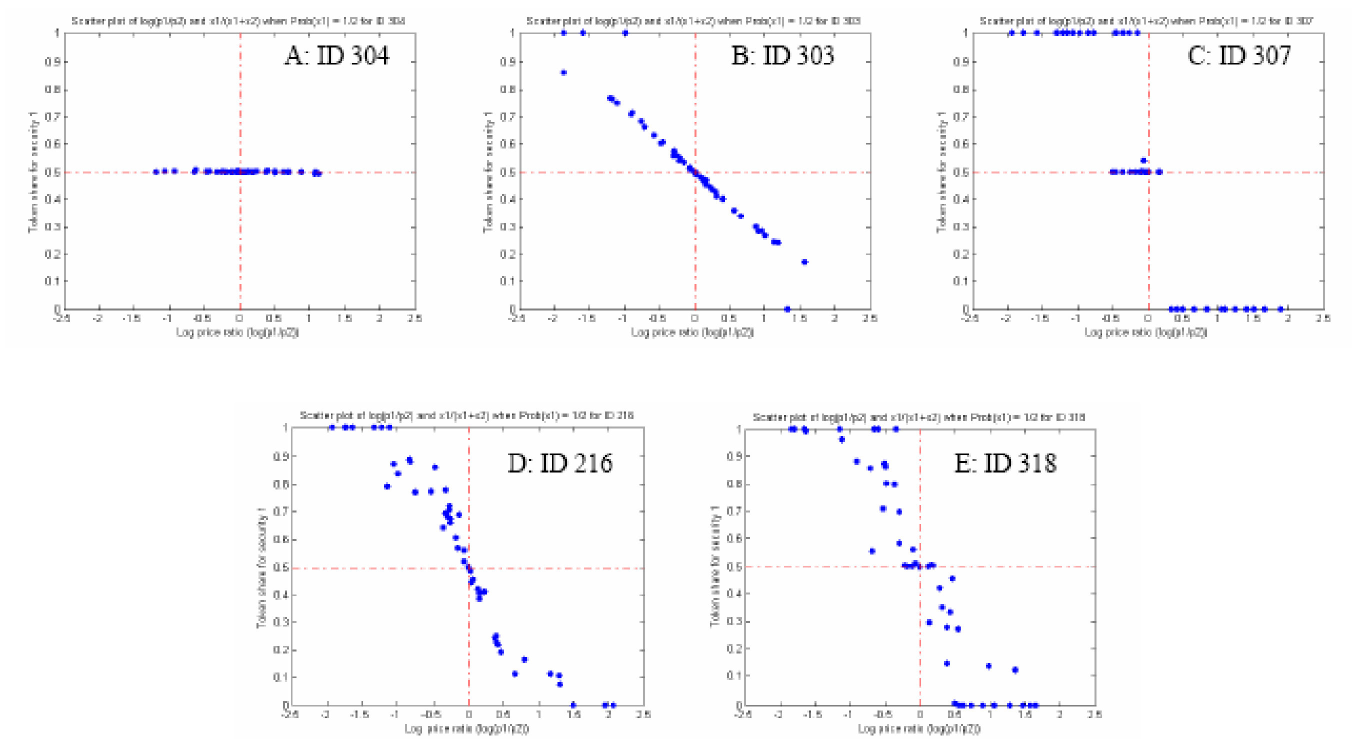
Proof is similar to proof of (i); (ii) is a consequence of state-independence and the separability across states implied by the independence axiom.

(iii) Risk aversion makes the indifference map's "better than" sets convex, because $p_1u(x_1) + p_2u(x_2)$ is a concave function of (x_1, x_2).

Now look at some subjects from Choi, Fisman, Gale and Kariv, "Consistency and Heterogeneity of Individual Behavior under Uncertainty" *AER* December 2007 or at http://socrates.berkeley.edu/~kariv/CFGK_III.pdf:

Figure 2 graphs $x_1/(x_1+x_2)$ against $\ln(p_1/p_2)$ for subjects who faced many choices like in Figure 1.

Figure 2: The relationship between the log-price ratio $\ln(p_1/p_2)$ and the token share $x_1/(x_1+x_2)$ for selected subjects.



What do the indifference maps look like for a subject who "bunches" at different price ratios, like subject 304? For a subject who either bunches or flips to extreme choices for extreme prices, like subject 307? Can you get such indifference maps with expected-utility preferences?

First-order risk aversion and reference-dependent preferences

Like the more direct evidence from Survey questions 1 (a and b) and 2 (a and b), and the Mugs experiment, the widespread "first-order" aversion to even small or moderate risks revealed by decisions to turn down moderate gambles with positive expected returns, or the bunching in Choi et al.'s experiments, suggests an explanation via reference-dependent preferences.

Rabin (2000 *Econometrica*) and Rabin and Thaler (2001 *JEP*) close off most escape routes:

● Ambiguity aversion doesn't help, because the reactions that cause the problem are to known probabilities, hence separate from those that underlie the Ellsberg paradox.

● Allowing preferences over final wealth distributions that are nonlinear in the probabilities doesn't help, because Safra and Segal, "Calibration Results for Non-Expected Utility Theories" http://fmwww.bc.edu/EC-P/WP645.pdf, show that the reactions are separate from those that underlie the Allais paradox.

● And—a technical point—nondifferentiable kinks can't be ubiquitous enough to save the vN-M theory from the Choi et al. bunching because the typical reactions to risks hold everywhere but a concave von Neumann-Morgenstern utility function must be differentiable almost everywhere.

This means that if we accept the notion of a preference-based model of choice under uncertainty, the explanation of first-order risk aversion must involve relaxing some other assumption.

Kahneman and Tversky (1979 *Econometrica*) in their Prospect Theory, Koszegi and Rabin (*QJE* 2006, *AER* 2007), and others who have considered this question favor relaxing the assumption that the “outcomes” over which utility functions are defined are lifetime final wealth levels.

They propose to replace this conventional assumption by the alternative that preferences are defined over gains and losses in wealth measured relative to some “reference point.”

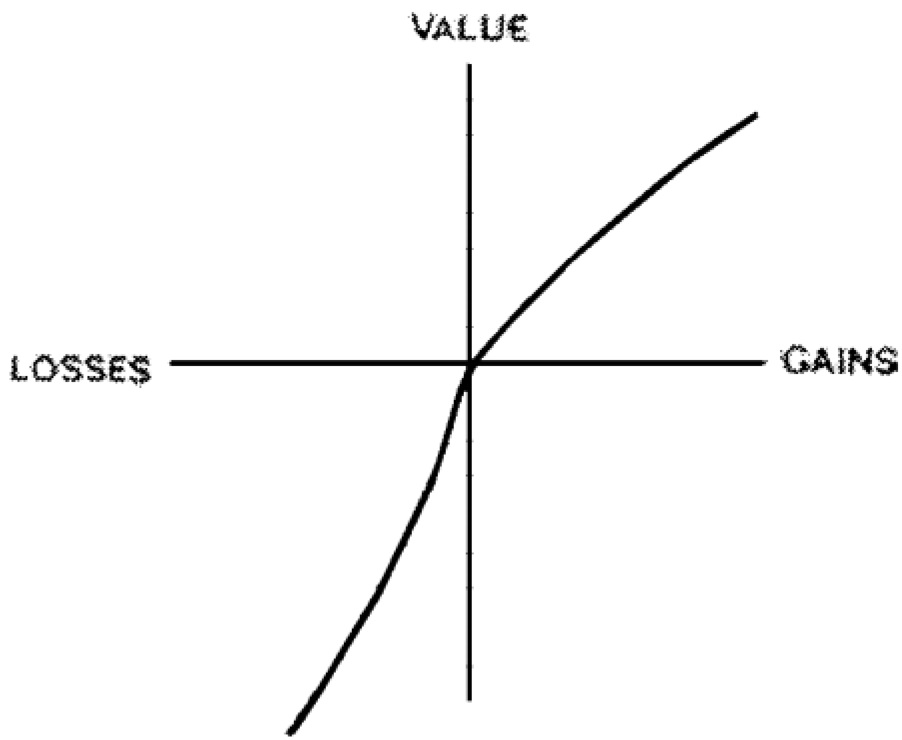


FIGURE 6. A SCHEMATIC VALUE FUNCTION FOR CHANGES

From Kahneman (December 2003 *AER*)

Kahneman and Tversky (1979 *Econometrica*, p. 277) stress that the salience of changes from reference points in their Prospect Theory is a basic aspect of human nature:

> An essential feature of the present theory is that the carriers of value are changes in wealth or welfare, rather than final states. This assumption is compatible with basic principles of perception and judgment. Our perceptual apparatus is attuned to the evaluation of changes or differences rather than to the evaluation of absolute magnitudes. When we respond to attributes such as brightness, loudness, or temperature, the past and present context of experience defines an adaptation level, or reference point, and stimuli are perceived in relation to this reference point (Helson (1964)). Thus, an object at a given temperature may be experienced as hot or cold to the touch depending on the temperature to which one has adapted. The same principle applies to non-sensory attributes such as health, prestige, and wealth. The same level of wealth, for example, may imply abject poverty for one person and great riches for another depending on their current assets.

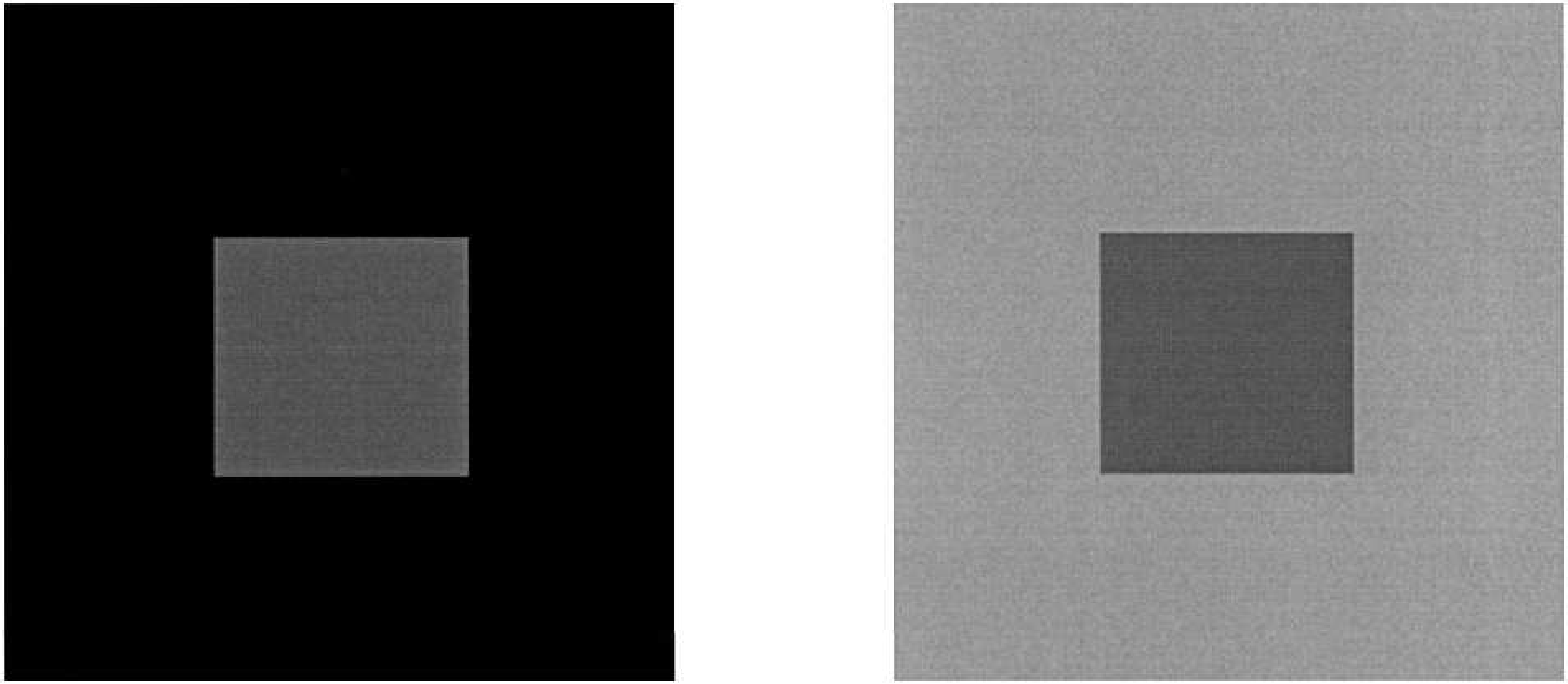


FIGURE 5. REFERENCE-DEPENDENCE IN THE PERCEPTION OF BRIGHTNESS

From Kahneman (December 2003 *AER*)

(The two inner squares are equally bright.)

FIGURE 7. AN ILLUSION OF ATTRIBUTE SUBSTITUTION

From Kahneman (December 2003 *AER*)

(The two horses are exactly the same size. Go ahead, measure them!)

Main features of reference-dependent preferences

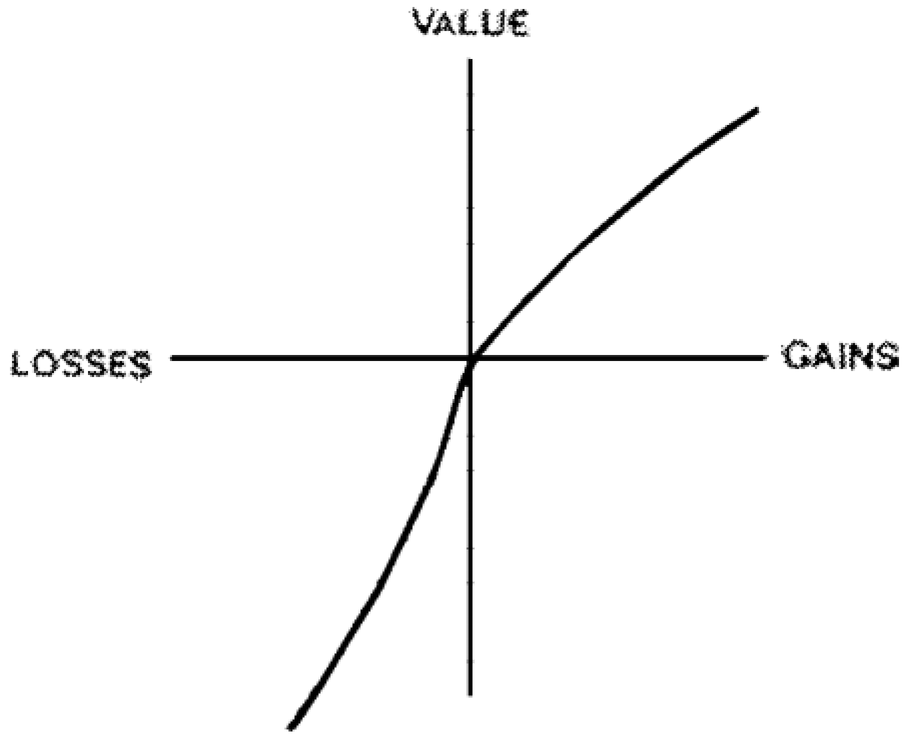


FIGURE 6. A SCHEMATIC VALUE FUNCTION FOR CHANGES

● **Loss aversion**

Note the kink at 0 (which is taken to represent the reference point here), which means that a small decrease (in income or wealth) below the reference point hurts (in value) more than an equally small increase above the reference point helps.

(The "coefficient of loss aversion" is defined as the ratio of marginal value loss below to marginal value gain above the reference point; when measured it is usually about 2 or 2½.)

● **Diminishing sensitivity (to losses as well as gains)**

Kahneman and Tversky (1979) (hypothetical questions): Which would you prefer?

0.45 chance of gaining \$6000 vs. 0.90 chance of gaining \$3,000: 14% chose 0.45 chance of \$6000.
0.45 chance of losing \$6000 vs. 0.90 chance of losing \$3,000: 92% chose 0.45 chance of \$6,000.

Or recall Survey questions 1 (a and b):

1a. Would you choose to lose \$500 for sure or to lose \$1000 with probability 0.5?
1b. Would you choose to receive \$500 for sure or to receive \$1000 with probability 0.5?

Most people give "risk-loving" answers to 1a (lose \$1000 with probability 0.5) but "risk-averse" answers to 1b (receive \$500 for sure).

As with reference-dependence, Kahneman and Tversky (1979) argue that diminishing sensitivity reflects a more fundamental feature of human cognition and motivation:

> Many sensory and perceptual dimensions share the property that the psychological response is a concave function of the magnitude of physical change. For example, it is easier to discriminate between a change of 3 and a change of 6 in room temperature, than it is to discriminate between a change of 13 and a change of 16. We propose that this principle applies in particular to the evaluation of monetary changes…. Thus, we hypothesize that the value function for changes of wealth is normally concave above the reference point ... and often convex below it....

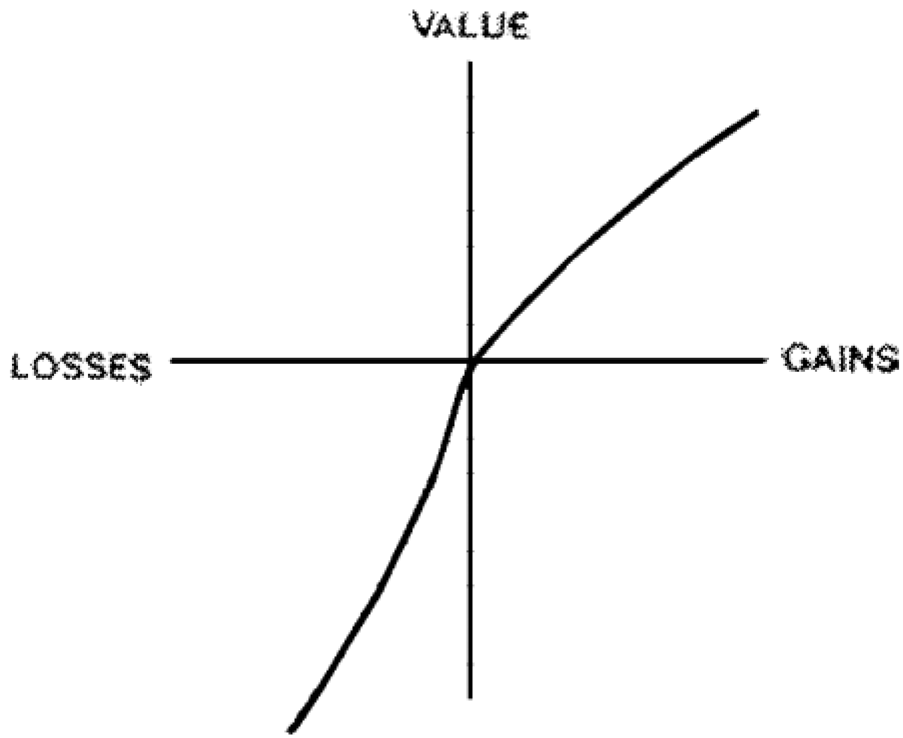


FIGURE 6. A SCHEMATIC VALUE FUNCTION FOR CHANGES

Diminishing marginal sensitivity to losses as well as gains: value function concave for $x > 0$, value convex for $x < 0$.

● **Nonlinear probability weighting (can't be seen in the picture!)**

A third feature of Prospect Theory, nonlinear probability weighting, by which people tend to overweight small probabilities and underweight large ones relative to expected utility theory (so the value of a risk is $\pi(p)v(x) + \pi(q)v(y)$, not $pv(x) + qv(y)$), is realistic and important, but less important than loss aversion and diminishing sensitivity, and will not be discussed here.

Re-doing standard economic analyses with Prospect Theory preferences

Although the literature sometimes makes a big deal about Prospect Theory's features of diminishing sensitivity and nonlinear probability weighting, the major difference between Prospect Theory and Expected Utility Theory is reference-dependence with loss aversion.

From now on I will treat "Prospect Theory" as synonymous with reference-dependence with loss aversion.

If, as is done here, we agree to ignore nonlinear probability weighting, we can re-run the von Neumann-Morgenstern Theorem to justify expected (Prospect Theory) value maximization.

The key point is that although expected utility is conventionally defined over lifetime final wealth levels, the logic of the von Neumann-Morgenstern Theorem works for preferences defined over anything.

Thus we are free to take preferences to be over gains or losses relative to a reference point, and the Theorem will ensure the existence of a (Prospect Theory) value function such that the person whose preferences are represented acts as if to maximize its expectation.

Most of the action in Prospect Theory comes just from reference-dependence and loss aversion relative to gains and losses from the reference point.

Although diminishing sensitivity and nonlinear probability weighting are realistic and important for some applications, it's possible to do a lot without them, using a piecewise linear value function with a coefficient of loss aversion of approximately 2.

(*Nobody* has a coefficient of loss aversion less than 1. The coefficient does seem to vary a bit from person to person, and perhaps from context to context (is it more painful to lose an apple or a banana? On Tuesday or Friday? etc.), but it's remarkably stable for an empirical parameter.)

With a tractable model of the reference point and a reasonable parametric specification of diminishing sensitivity (more or less like a vN-M utility function, but allowing a flip from convex to concave at the origin), Prospect Theory is still a bit less tractable than Expected Utility Theory, but not impossibly so.

With a piecewise linear value function and a simple model of the reference point, Prospect Theory may even be more tractable than Expected Utility Theory.

Below I give several example applications of this kind of Prospect Theory model.

Prospect Theory preferences and "first-order" risk aversion

Prospect Theory preferences with loss aversion have an "automatic," portable kink at the reference point which allows the theory easily to accommodate first-order risk-aversion, even with an otherwise differentiable value function.

Even a piecewise linear Prospect Theory value function with a coefficient of loss aversion of approximately 2 can explain first-order risk aversion.

Suppose that the value function is linear except for a kink at the reference point. If the value function is $v(\sigma)$, the reference point is normalized to 0 with $v(0) = 0$, and the coefficient of loss aversion is 2, then the value function is $v(\sigma) \equiv \sigma$, $\sigma > 0$, and $v(\sigma) \equiv 2\sigma$, $\sigma < 0$.

A person with such a value function will take a 50-50 win σ-lose σ gamble with risk premium π if and only if $\frac{1}{2}(\sigma + \pi) + \frac{1}{2}(-2\sigma + \pi) \geq 0$, which is true if and only if $\pi \geq \sigma/2$.

This is known as "first-order" risk aversion because the required risk premium grows linearly with the scale of the bet σ (note that the "scale" here is the same as the standard deviation).

(If we did this with a nonlinear value function using Taylor's Theorem, we'd get a similar formula for small-scale bets, in which the coefficient of loss aversion is defined as the ratio of the limiting marginal values for gains and losses approaching 0.)

Aside: "Second-order" risk aversion and approximate risk-neutrality of differentiable expected-utility maximizers over small bets

vN-M expected-utility maximizers with differentiable utility functions have only "second-order" risk aversion, in that the utility loss from a small gamble is proportional to its variance (rather than its standard deviation as with loss aversion, which would be "first-order").

Proof: A person with vN-M utility function u( ) and base wealth x will take a 50-50 win σ-lose σ gamble with risk premium π if and only if ½ $u(x + \sigma + \pi)$ + ½ $u(x - \sigma + \pi) \geq u(x)$.

Expanding the left-hand side in a Taylor Series around $x + \pi$, $u(x + \pi) + ½\, u'(x + \pi) - ½\, u'(x + \pi) + 2(½)^2 u''(x + \pi)\sigma^2 = u(x + \pi) + ½u''(x + \pi)\sigma^2 \geq u(x)$.

Expanding now in π about x, neglecting the small u''' term, and solving yields $u(x) + \pi u'(x) + ½u''(x)\sigma^2 \geq u(x)$ if and only if $\pi \geq -½[u''(x)/u'(x)]\sigma^2$, so that the required risk premium grows with σ^2: second-order risk aversion.

(Here, and π is proportional to the utility loss, and the factor $-[u''(x)/u'(x)] > 0$ is the Arrow-Pratt coefficient of absolute risk aversion, which was 1 for the Prospect Theory value function.)

“Second-order” risk aversion and insanity

With a differentiable, increasing von Neumann-Morgenstern utility function defined over money outcomes, a risk-loving person (convex utility function) must take any bet that is “more than fair” (strictly positive expected return).

With a differentiable, increasing von Neumann-Morgenstern utility function defined over money outcomes, a risk-averse person may turn down some more than fair bets, because the “cost” of a large risk may outweigh the positive expected return.

But if such a differentiable, risk-averse von Neumann-Morgenstern person is offered a more than fair bet with the option to scale it down as much as desired (e.g. changing a 50-50 win $11,000-lose $10,000 bet to a 50-50 win $1100-lose $1000 bet or, if he’s a total wimp, to a 50-50 win $110-lose $100 bet), then he must always take the bet at some strictly positive scale.

In effect, all differentiable vN-M people become approximately risk-neutral for small bets.

This is why people turning down small bets in Rabin and Thaler’s examples, together with the assumption that they have globally risk averse von Neumann-Morgenstern utility functions over final wealth, implies that they will be insanely risk-averse over large more-than-fair bets.

Application: Return of the Mug People

Recall that Kahneman, Knetsch, and Thaler (1990, 1991) randomly gave mugs to half the subjects in a classroom experiment ("owners") and nothing to the others ("non-owners").

They then elicited selling prices for owners and buying prices for non-owners. Supply and demand "should" be mirror images of each other. But…

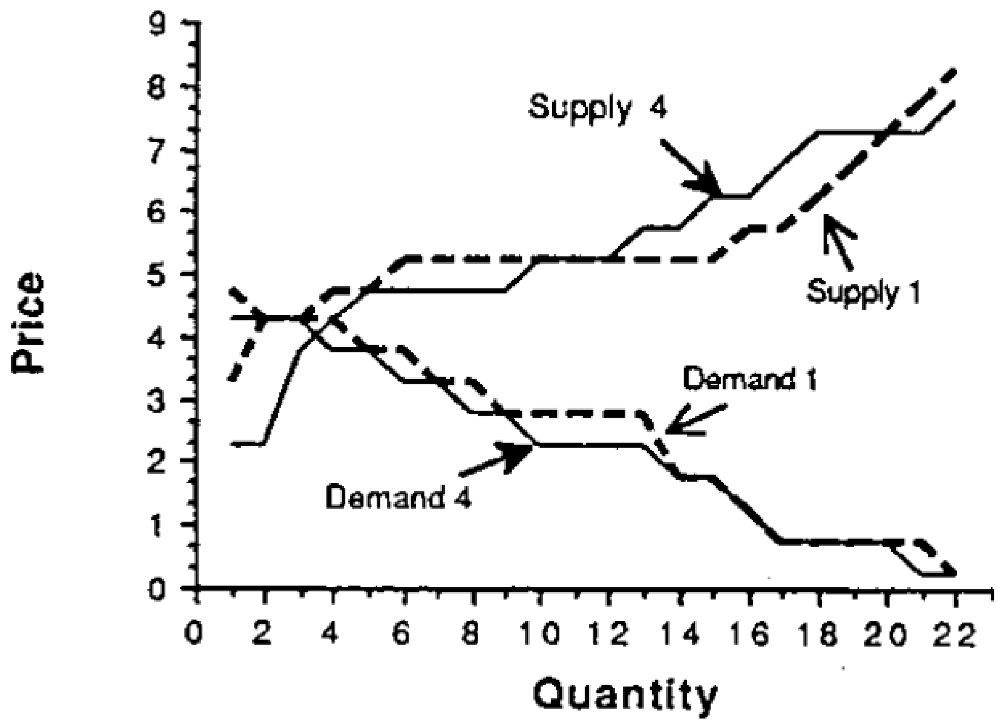


FIG. 1.—Supply and demand curves, markets 1 and 4

The average buying price of non-owners was about $3.50, and the average selling price of owners was about $7.00: way too big a gap to be random.

How do we model this with Prospect Theory's reference-dependence and loss aversion?

Imagine that (unlike Kahneman and Tversky, but like some of Koszegi and Rabin's more recent work) people have both ordinary consumption utilities for mugs and money, and gain-loss utilities (which Kahneman and Tversky focused on to the exclusion of consumption utilities, as may be approximately appropriate for laboratory experiments with small gifts).

Assume that subjects' (owners' and non-owners') consumption utilities for mugs are uniformly distributed between $0 and $9. They have linear consumption utility: value = value of mug (or not) + $.

Assume that subjects also have gain-loss utilities, with no diminishing sensitivity but with a coefficient of loss aversion of 2, so that losses relative to the reference point lower their gain-loss utility twice as much as gains raise it.

The weight of gain-loss utility is η, so total utility is consumption utility + $\eta \times$gain-loss utility.

Subjects' reference points are determined by their expectations:

Owners expect to keep their mugs (and gain no money).

Non-owners expect to keep their money (and gain no mug).

Supply of mugs

An owner with mug consumption value $$v$ who is considering trading her/his mug for $$m$ will compare his total (consumption plus gain-loss) utility from keeping her/his mug with her/his total utility from trading the mug for $$m$.

Because as an owner s/he expected to keep her/his mug, if s/he keeps it there are no gain-loss surprises on either the mug or the money dimension.

Her/his total utility from keeping = consumption utility ($v + 0$) + $\eta \times$gain-loss utility ($0 + 0$).

If s/he trades her/his mug for $$m$, there are gain-loss surprises on both dimensions, "losing" her/him $\eta \times 2v$ on the mug dimension—because it's *her/his* mug, and the coefficient of loss aversion is 2—but gaining her/him $\eta \times m$ on the money dimension—only m, because it's *someone else*'s money.

Her/his total utility from trading = consumption utility ($0 + m$) + $\eta \times$gain-loss utility ($-2v + m$).

Thus the lowest price m at which s/he would be willing to sell her/his mug is the lowest m that makes $v \leq m + \eta(-2v + m)$, or $m^* = v(1 + 2\eta)/(1 + \eta)$.

If $\eta = 0$ we get the usual $m^* = v$ result; but if $\eta > 0$, say $\eta = 1$, we get $m^* = 1.5v$, which yields an average selling price of $6.75 ≈ Kahneman, Knetsch, and Thaler's $7.

(If you consider the whole distribution of values, it's easy to generate a supply curve as above.)

Demand for mugs

Similarly, a non-owner with mug consumption value \$$v$ who is considering trading \$$m$ of her/his (hard-earned!) money for a mug will compare her/his total (consumption plus gain-loss) utility from keeping her/his \$$m$ with her/his total utility from trading \$$m$ for a mug.

Because as a non-owner s/he expected to keep her/his \$$m$, if s/he keeps it there are no gain-loss surprises on either the money or the mug dimension.

Her/his total utility from keeping = consumption utility $(0 + m)$ + $\eta\times$gain-loss utility $(0 + 0)$.

If s/he trades her/his \$$m$ for a mug, there are gain-loss surprises on both dimensions, gaining her/him $\eta\times v$ on the mug dimension but losing her/him $\eta\times 2m$ on the money dimension.

Her/his total utility from trading = consumption utility $(v + 0)$ + $\eta\times$gain-loss utility $(v\text{-}2m)$.

Thus the highest price m^ s/he would be willing to pay for the mug is the highest m that makes $v + \eta(v - 2m) \geq m$, or m^ $= v(1 + \eta)/(1 + 2\eta)$.

If $\eta = 0$ we get the usual m^ $= v$ result; but if $\eta > 0$, say $\eta = 1$, we get m^ $= 0.67v$, which yields an average buying price of \$3.00 ≈ Kahneman, Knetsch, and Thaler's \$3.50.

(If you consider the whole distribution of values it's easy to generate a demand curve as above.)

You should re-do the above argument, with $\eta = 1$, for a mug-owner who expects to sell her/his mug, say for $\$x$ (so her/his reference point is having $\$x$ and no mug). Then re-do it for a non-owner who expects to buy a mug for $\$y$ (so her/his reference point is having a mug but $-\$y$).

You will find that these expectations make both sellers and buyers more willing to trade. Expectations create a preference bias, relative to the standard model, in favor of what was expected. This is the reasoning behind this quotation from Koszegi-Rabin (2006 *QJE*):

> …when expectations and the status quo are different—a common situation in economic environments—equating the reference point with expectations generally makes better predictions. Our theory, for instance, supports the common view that the "endowment effect" found in the laboratory, whereby random owners value an object more than nonowners, is due to loss aversion—since an owner's loss of the object looms larger than a nonowner's gain of the object. But our theory makes the less common prediction that the endowment effect among such owners and nonowners with no predisposition to trade will disappear among sellers and buyers in real-world markets who expect to trade. Merchants do not assess intended sales as loss of inventory, but do assess failed sales as loss of money; buyers do not assess intended expenditures as losses, but do assess failures to carry out intended purchases or paying more than expected as losses.

The non-owner's decision is just like the one in the "shopping for shoes" example from Koszegi and Rabin's (2006 *QJE*) paper, which makes some interesting (though more difficult) further points. See also problem 22 on Problem Set 1. In Koszegi and Rabin's examples with price uncertainty the calculations get harder but the basic ideas are the same as above.

More applications

Let's start with the applications I referred to in the Introduction:

● The phenomenon that race-track bettors tend to bet more on long shots near the end of the day at the track.

Here it's natural to take the reference point as breaking even and the period over which gains and losses are evaluated (the "bracket") as the day at the track. Loss aversion without diminishing sensitivity is enough to generate betting strategies that vary with gains or losses during the day in a way that makes long shots look more attractive to most bettors (losers) near the end of the day. At the end of the day, most people have lost money, and so are willing to take risks to break even. Hence it makes sense for them to bet on (risky) long shots, because a small bet placed on a long shot can generate enough profit to cover the day's losses. The effect is strong enough to make betting on the favorite to show in the last race profitable.

● The phenomenon that house sellers who paid more for their houses (a sunk cost in standard theory) set asking prices that are higher, controlling for quality, so that they tend to take longer to find a buyer, but to sell at a higher price.

Here it's natural to take the reference point as breaking even relative to what you paid for the house (apparently without controlling for inflation), and the natural bracket is the purchase and sale of a given house (i.e. you don't mentally trade off losses on one house against gains on another, or gains from selling your Ferrari).

Genesove and Mayer (2001 *QJE*) studied the market for Boston condominiums sold between 1990 and 1997 by sellers who originally purchased the houses after 1982. If person A bought a condo of a given quality at a 10% higher price than person B bought a similar condo, because she bought at a time when the market averaged 10% higher, new buyers will value the houses equally. In the absence of other differences A and B should have the same selling prices.

In the data, however, there are dramatic differences: Sellers who are selling their condos for a loss (in nominal terms) relative to their buying price charge a higher price than those selling equally-valued homes without a loss—by on average 35% of the average difference between the appropriate price and the price at which they bought it.

Say two people each have a house currently valued at $500,000, but A bought it for $600,000 while B bought it for $500,000. Then A will ask $535,000 and B will ask $500,000.

Genesove and Mayer carefully rule out other possible explanations, leaving loss aversion. Investor sellers exhibit less loss aversion than owner-occupier sellers, but still have some.

Now for some new applications:

● The phenomenon that investors in the stock market are reluctant to realize losses.

Odean (1998 *Journal of Finance*) finds that small investors (without brokers) are much more likely to sell winners than losers.

If, as for houses, it's natural to take the reference point as breaking even relative to what you paid for the stock, loss-averse people who have lost money on the stock will tend (with diminishing sensitivity) to be risk-loving, willing to take risks and wait for the price to recover before selling. By contrast, people who have made money on the stock will tend to be risk-averse, hence more likely to sell.

Odean also checked alternative explanations:

The winners that small investors sell do better than the losers they hold, so it's unlikely that they are extrapolating expectations from past performance.

Tax considerations should lead investors to sell losers rather than winners, to decrease current taxes, so it's unlikely that they are doing it to save on taxes.

● "Deal or No Deal?"

Post, van den Assem, Baltussen, and Thaler, "Deal or No Deal? Decision Making Under Risk in a Large-Payoff Game Show" (2008 March *AER*; now on the *AER* web site as a forthcoming paper) study European versions of the game show "Deal or No Deal," in which contestants make a sequence of risky decisions with huge stakes.

In the game, a contestant "owns" a suitcase with a randomly determined prize. Gradually, the contestant learns information about the prize in her bag (by opening other bags and learning what is not in her bag). At each stage, a "bank" offers a riskless amount of money to replace the amount in the bag. A contestant's acceptance or rejection of the offer can be used to infer her/his risk aversion.

The authors find strong evidence of reference-dependence, in that contestants become more risk-accepting when they have received bad news in the last few rounds.

● The phenomenon that people sometimes pay huge premia to insure against trivial risks.

A leading example is Cicchetti and Dubin's (1994 *JPE*) analysis of people's decisions of whether to buy insurance against damage to their home telephone wiring, where they found that people would pay almost twice the expected cost to insure against a loss of less than $100.

Another, though less trivial example is Justin Sydnor, "Abundant Aversion to Moderate Risk: Evidence from Homeowners Insurance," 2006, http://wsomfaculty.case.edu/sydnor/deductibles.pdf , who studied people's choice of deductible for home insurance.

In his sample, customers can choose between four deductible levels: $100, $250, $500, and $1,000. Because his data include house characteristics, he not only knows the deductible people chose, how much premium they are paying, and the claims they made, but how much they would have paid and/or received had they chosen a different deductible.

Almost nobody chooses the $100 deductible, but the other deductible levels are chosen by a large number of people. People overpay for lower deductibles by a factor of 5.

● Insurance even for your ferret or your tea kettle?!? Coming soon to America…

Rabin asks: Why can you buy an extended warranty on your tea kettle in England or can you insure your ferret in Sweden, and why do companies work so hard to sell you such insurance?

● What? Ferrets and tea kettles are trivial examples? Okay, how about the entire stock or labor market? Camerer et al. (1997; the cab driver paper discussed below) give a good summary of an application of loss aversion to the famous "equity premium puzzle":

> Benartzi and Thaler [1995] use the same combination of narrow bracketing and loss aversion that we use, to explain the equity premium puzzle—the tendency for stocks to offer much higher rates of returns than bonds over almost any moderately long time interval. In their model, the equity premium compensates stockholders for the risk of suffering a loss over a short horizon. They show that if investors evaluate the returns on their portfolios once a year (taking a narrow horizon), and have a piecewise-linear utility function which is twice as steep for losses as for gains, then investors will be roughly indifferent between stocks and bonds, which justifies the large difference in expected returns. If investors took a longer horizon, or cared less about losses, they would demand a smaller equity premium. Two papers in this issue [Thaler, Tversky, Kahneman, and Schwartz 1997; Gneezy and Potters 1997] demonstrate the same effect in experiments.

The argument is similar to the one I gave regarding loss aversion and first-order risk aversion.

Published references:

Benartzi and Thaler, "Myopic Loss Aversion and the Equity Premium Puzzle" (1995 *QJE*).
Siegel and Thaler, "The Equity Premium Puzzle" (1997 *Journal of Economic Perspectives*).

● Labor supply

Camerer, Babcock, Loewenstein, and Thaler, “Labor Supply of New York City Cabdrivers: One Day at a Time” (1997 *QJE*) study a potentially important and influential application.

Cab drivers are great for testing theories of intertemporal labor supply because unlike most workers they choose their own hours each day, and conditions are roughly constant within a day.

Theories of labor supply play an important role in labor economics and macroeconomics, where they have a major impact on the interpretation of business cycles and assessment of their costs.)

Standard choice theories all predict a positive relationship between daily wages and hours worked—intertemporal substitution—because income effect of a change in daily wage is negligible.

But correlations between log hours and log wages are strongly negative, between -0.503 and -0.269, with elasticities close to -1 for experienced drivers:

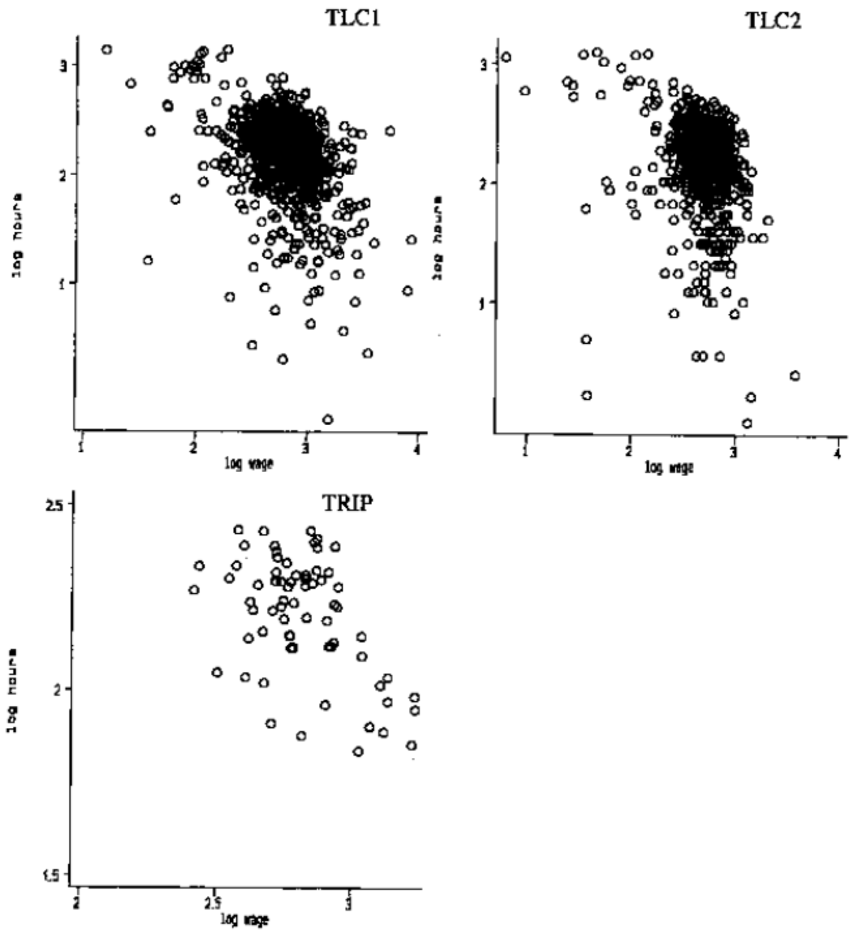


FIGURE I
Hours-Wage Relationships

The elasticities are as if drivers had a daily income target (narrow bracketing) and worked until they reached it.

Note how this reduces earnings: if you reach the target very early, it's a signal that you could earn a lot more by working longer that day.

The authors (see also Koszegi and Rabin (*QJE* 2006)) propose an explanation in terms of reference-dependent preferences via daily income targeting.

Here, the bracket is the day, and the target is presumably set by past experience in some way (Koszegi and Rabin propose models).

Falling short of the day's target is a painful loss, while going above it is less rewarding than in standard theories, relative to the costs: so there's a kink at the target, whatever it is.

Daily income targeting easily explains the negative correlation between wages and hours.

The authors carefully checked alternative explanations.

See also Henry Farber, "Is Tomorrow Another Day? The Labor Supply of New York City Cabdrivers," 2005 *Journal of Political Economy*

Farber, "Reference-Dependent Preferences and Labor Supply: The Labor Supply of New York City Cab Drivers," (2008 *AER*)

Juanjuan Meng, "Are Hours and Income Both Targeted? A Multi-targeting Model of New York Cabdrivers' Labor Supply Behavior," UCSD

Unresolved issues: "Mental accounting" and "narrow bracketing"

Mental accounting and narrow bracketing are two important issues given short schrift here.

Note that having a reference point for anything less than everything that happens to you in your lifetime logically requires a theory of "mental accounting" with "narrow bracketing":

- What gains/losses are grouped together?

- When are mental accounts closed/opened?

- How do time, space, and cognitive boundaries affect them?

Some answers to these questions are implicit in the applications discussed above. For example, the fact that race-track bettors' and cab drivers' behavior seems to be organized day by day suggests that they have daily mental accounts. (If their behavior had seemed to change between mornings and afternoons, and according to cumulative morning or afternoon totals over the week, we would need a more complex notion of mental accounts to define loss aversion.) By contrast, Benartzi and Thaler's explanation of the equity premium puzzle assumes that investors evaluate their positions year by year. Both specifications are plausible for their applications, but we have as yet no theory that determines them. The questions are empirical—about behavior, not logic—but fortunately there are empirical regularities to guide assumptions about them.

A good place to start reading about this is Thaler, "Anomalies: Savings, Fungibility, and Mental Accounts," 1990 *Journal of Economic Perspectives*.