Second-Year Advanced Microeconomics: Behavioural Economics

Behavioural Decision Theory: Choice under Uncertainty and Certainty, Hilary Term 2010

Vincent P. Crawford, University of Oxford
(with very large debts to Botond Kőszegi and Matthew Rabin)

Last revised 23 December 2009

Introduction

I start with an extended quotation from Matthew Rabin's unpublished notes, summarizing the "behavioural" point of view (see also his 2002 *European Economic Review* paper):

Premise 1: Most (not all) facts about people that seem to be true and seem to be economically relevant–even if these assumptions didn't appear in the prior generation of economics textbooks are ... both true and economically relevant.

Premise 2: "Untraditional" or unfamiliar assumptions, including those that imply various limits to rational utility-maximization, can and should mostly be studied using exactly the same set of tools and approaches economists are used to (i.e., formal mathematical models and statistical tests using laboratory and especially field data), using exactly the same scientific criteria (good predictions, parsimony, etc.) as economists are used to. The sole difference in methods and goals of most economists is the broader array of aspects of human nature we study.

Premise 3: Not only are familiar economic methods great, but to a very large extent so are familiar economic assumptions. The fact that there are limits to the correctness and applicability of these assumptions does not mean that they aren't often exactly the appropriate assumptions—nor that they aren't tremendously useful even when not exactly right. The material in this course is not meant as a replacement of, but as an enhancement of—and eventual component of—mainstream economics…. The eventual goal … is that … "behavioral economics" will eventually disappear as a separate or isolated field….

(Still Rabin:)

Focusing for now on individual decisions, what does mainstream (standard neoclassical) economics normally assume?

Homo Economicus (possibly unlike Homo Sapiens):

● Is perfectly rational, making choices that consistently maximize some exogenous, stable set of preferences that depend on absolute levels of outcomes (rather than changes), even with uncertainty (in which case preferences are expected utility), and even in dynamic situations (in which case preferences are discounted sums of per-period expected utilities)

● Is also perfectly rational in the sense of costlessly and correctly making logical nonprobabilistic inferences and applying the laws of probability to process information and make probabilistic judgments (Bayes' Rule, contingent reasoning, option value)

● Has perfect will-power and the ability to make and follow intertemporal plans (even contingent ones), with no conflict between the preferences of current and future selves

● Is almost always also assumed to be perfectly self-interested, caring exclusively about her/his own consumption, though this assumption is not essential to mainstream theory

(Still Rabin:)

Familiar as they are, nothing essential depends on whether these assumptions are considered “standard”.

Even rationality, though very important to the way economics is done, is not an essential axiom without which no coherent theory is possible.

What are “standard” assumptions is only a convention of the discipline, subject to change when different assumptions appear to be more useful.

In many settings the standard assumptions are reasonable stylizations of the “facts” that most of the people whose behavior we wish to analyze are usually self-interested and well-informed, with coherent goals, and reasonable skill at making plans to realize them.

The standard assumptions also embody a kind of methodological humility, preventing us from assuming we know more about people’s goals and possibilities than they do.

But in other, equally important settings, standard assumptions are unreasonable as descriptions of behavior.

This means more than that they are literally incorrect; all behavioral assumptions are incorrect to some degree. It means that they are systematically incorrect, and that the incorrectness has economically important consequences; otherwise we would be happy to stick with the standard model.

(Still Rabin:)

In such settings the standard assumptions have sometimes significantly hindered research.

They have allowed/encouraged economists to ignore research that directly explores preferences, beliefs, information processing, and other determinants of human behavior.

There was no need to study the structure of preferences directly, because under the standard assumption that choices maximize preferences, preferences can be inferred from choices via revealed preference.

There was no need to study belief formation or information processing, because it was assumed to be completely determined by rationality postulates.

Behavioral economists, by contrast, have to be open to alternative ideas and evidence about preferences, beliefs, and information processing.

In such settings, behavioral economists (unlike some mainstream economists) are willing to consider deviations from standard assumptions in directions suggested by behavioral evidence, if it yields better explanations of outcomes than standard models do.

To paraphrase Rabin in a sound bite, behavioural economics *is* neoclassical economics, but taking a softer and less tradition-dominated line on behavioural assumptions.

The prototypical economist's conception of human behavior is roughly that people choose among all feasible lifetime plans to maximize a lifetime sum of discounted, additively separable expected utilities, with beliefs that are formed in a rational, Bayesian way.

Ancillary assumptions are often added, such as self-interest, non-habituation, convexity and smoothness of preferences.

There are four leading directions in which the standard model of individual decisions might be improved by making behaviourally more realistic assumptions:

- Choice under uncertainty or certainty

- Probabilistic judgment

- Present-biased preferences and time-inconsistency in intertemporal choice

- Social preferences, including altruism, envy, spite, and reciprocity.

These lectures will focus on the first two directions (more may be added in future years).

Review of neoclassical rational choice theory: Representation of preferences over certain outcomes via utility maximization

I describe the theory as if it were static, but my description it applies equally well to dynamic choice with the objects of choice viewed as dynamic decision rules rather than static decisions.

For choice under certainty, if preferences are complete, transitive, and continuous, then one can construct a utility function (ordinal) such that the individual chooses "as if" to maximize utility.

(Continuity is needed only if choices are continuously variable, and then only to rule out technical difficulties such as those caused by lexicographic preferences.)

The utility function is just a compact, tractable way to describe consistent choices in various settings, allowing us to store intuition about behaviour from simple experiments or thought-experiments and transport it to new situations.

Rationality in this narrow sense of choice consistency is a fairly weak assumption.

Logically, as long as choices are consistent, the preferences that represent them can be anything: self-interested or not, increasing in intuitive directions (more income) or not, etc.

(From Amartya Sen's famous paper critiquing the common neoclassical assumption of narrow self-interest: "Rational Fools" (1977 *Philosophy and Public Affairs*): "But if you are consistent, then no matter whether you are a single-minded egoist or a raving altruist or a class conscious militant, you will appear to be maximizing your own utility in this enchanted world of definitions.")

The flip side of this is that, unless we commit to a particular specification of preferences, the theory of rational choice is flexible enough to allow (almost) anything, and is therefore (almost) useless.

Partly for this reason, mainstream neoclassical economics has very strong conventions about what assumptions about preferences are “reasonable”: e.g. that they respond only to own (rather than own and others’) income or consumption; to income or consumption without regard to how it is generated; and to levels of income or consumption rather than changes.

If you don’t agree with a neoclassical analysis of something, it might be because you don’t agree with these conventions as applied in that analysis, not because you don’t agree with rationality in the general sense of choice consistency.

Which is not to deny the existence of irrationality even in the general sense; just to deny that relaxing rationality is always necessary, or the most useful way, to explain violations of neoclassical predictions.

As we’ll see, behavioural decision theory retains a lot of rationality in the general sense of choice consistency.

Instead of abandoning choice consistency, behavioural decision theory seeks progress (mainly) by modifying neoclassical conventions on what preferences are about, in realistic directions.

It seeks to avoid the lack of parsimony of rational choice without commitment to particular specifications of preferences by insisting that any deviations from neoclassical conventions be firmly grounded in evidence about behaviour.

This motivates a sharp focus on particular, minimal classes of deviations from neoclassical assumptions about choice under uncertainty or certainty, probabilistic judgment, time-consisteny intertemporal choice, or self-interested preference maximization.

Choice under certainty: Mugs and the willingness to pay-willingness to accept gap

In a famous experiment, Kahneman, Knetsch, and Thaler (“KKT”; 1990 *Journal of Political Economy*) randomly gave mugs to half the subjects in a classroom experiment (“owners”) and nothing to the others (“nonowners”).

They then elicited selling prices for owners and buying prices for nonowners.

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 front of the room (in plain view of all).

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”.

The highest price at which a seller expresses a willingness to keep the good is taken as her/his “selling price”.

If KKT 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 nonowners, on average, just because they were the ones who chose to acquire them, or perhaps because they had learned to love the mugs they had acquired.

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

However, in the experiment, owners and nonowners 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.

Nonetheless…

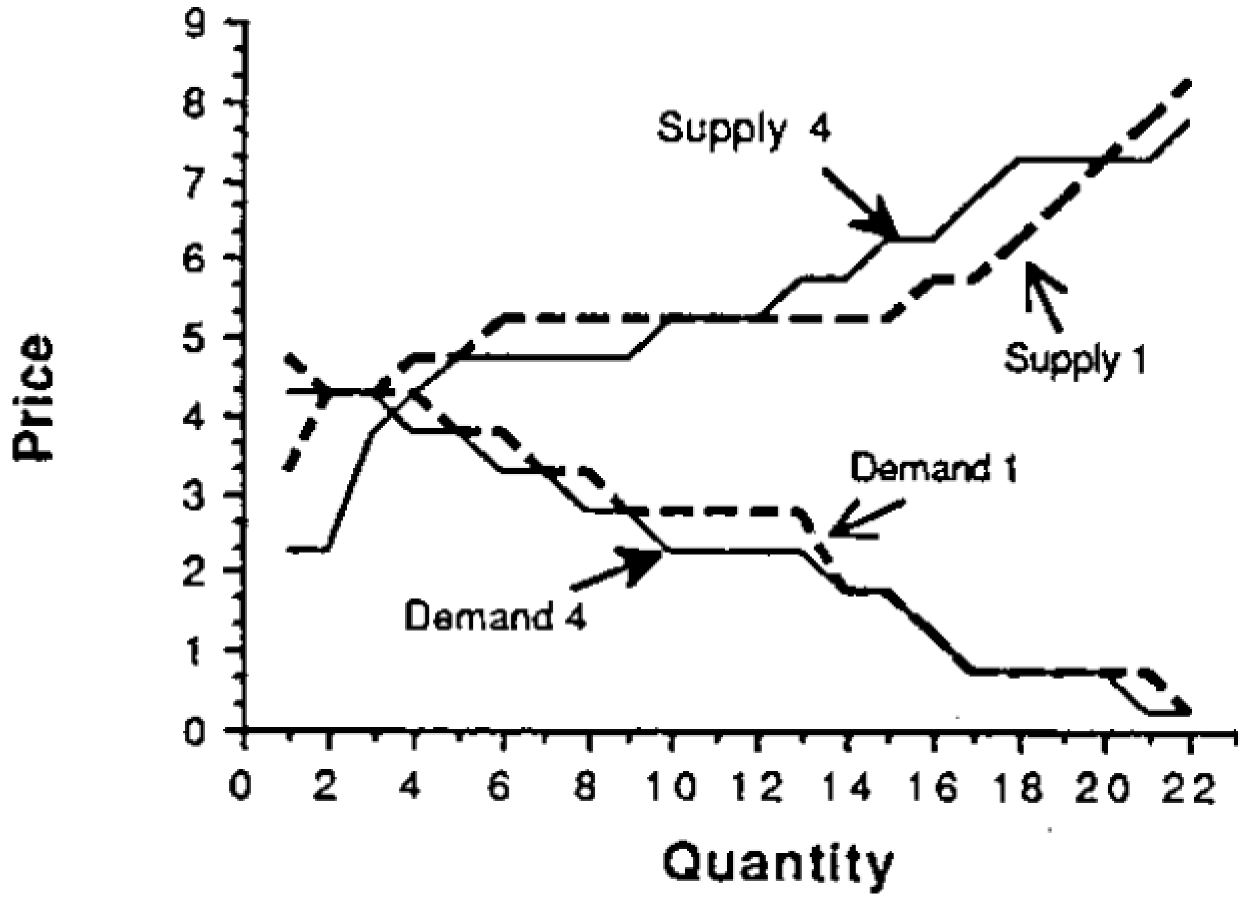

FIG. 1.—Supply and demand curves, markets 1 and 4

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

This result has been replicated many times, with the gap almost always in the same direction (see however the no-gap findings of Plott and Zeiler, 2005 and 2007 *AER*, and the forthcoming *AER* Comment by Isoni, Loomes, and Sugden and Reply by Plott and Zeiler).

Standard neoclassical demand theory would model this situation by postulating a utility function over *levels* of consumption of mugs and/or money (treating the latter as a proxy for other lifetime consumption).

In this framework, it is *logically* possible that gap is due to income effects.

For, those subjects who received mugs were, on average, slightly richer than those who did not.

And in theory, even a tiny change in income or wealth can radically alter a person's mug-money tradeoff.

But there’s no reason why people with varying initial wealths should all (or even most) have income effects that suddenly double (on average) their values for mugs, each at the status quo income level before the experiment.

In an econometric model of the demand for mugs, it would violate even neoclassical conventions to assume (or even allow) such a magical correlation between initial wealth and values for mugs.

Further, even if we allowed such a correlation, income effects from mugs are not large enough to plausibly explain a gap as big as $7.00 versus $3.50.

No reasonable specification of preferences *over levels alone* will make income effects into a credible explanation of the willingness to pay-willingness to accept gap.

Aside:

Income effects can be ruled out more definitively as an explanation by an experiment in which another group of subjects, “choosers,” are told they will be given either a mug or money, and asked to state the amount of money that makes them indifferent between the amount and the mug.

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

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.

End of aside

Prospect theory and reference-dependent preferences

A natural, behaviorally plausible, and parsimonious explanation of the gap, which is consistent with a large body of evidence from other settings, is to allow preferences to be reference-dependent—see especially Kahneman and Tversky's (1979 *Econometrica*) prospect theory and Kőszegi and Rabin's (2006 *QJE* 2006) generalization.

Reference-dependence expands the space over which preferences are defined to include a "reference point" as well as standard consumption, in this case of mugs and money.

Preferences are then defined, not over levels, but over *changes* in the sense of gains or losses relative to the reference point.

(As explained above, the assumption that preferences respond to levels rather than changes is not logically necessary to use utility maximization to describe choice.

It is only a convention of neoclassical economics, which could be replaced by an alternative convention if it were found to be useful.

Further, it may be more sensible to allow preferences to depend on both levels and changes. This is considered below, but for now let's focus on changes.)

In defining reference-dependent preferences to explain the gap, it seems natural to take the reference point as the status quo before the choice; but we will have to think harder about how to define the reference point in other applications.

It also seems natural to assume that subjects consider the mug-money choices in isolation, without trying to integrate them into a lifetime consumption plan; but we will also have to think harder about "mental accounting" and "bracketing": how people group choices in thinking about them.

Kahneman and Tversky (1979 *Econometrica*) stress that prospect theory's salience of changes from reference points 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.

Two examples from Kahneman (December 2003 *AER*):

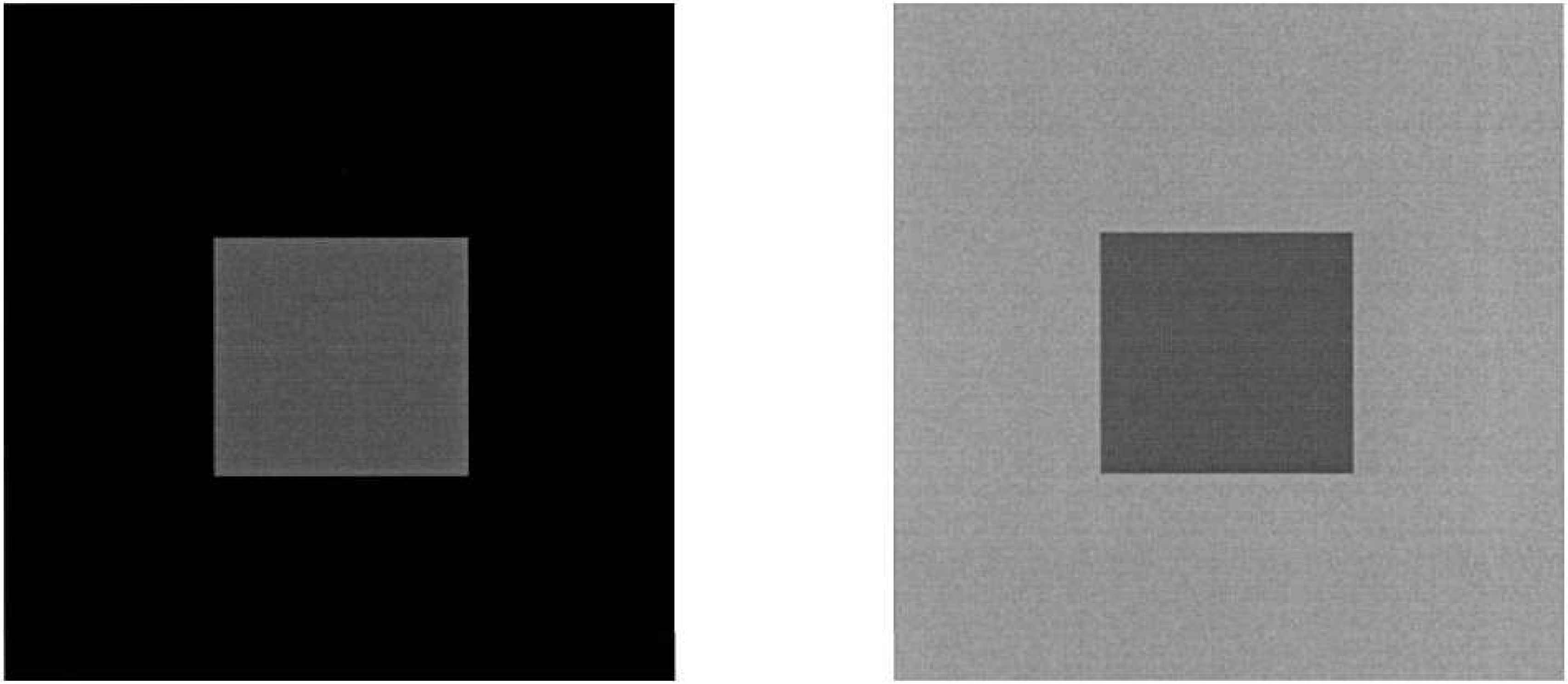

FIGURE 5. REFERENCE-DEPENDENCE IN THE PERCEPTION OF BRIGHTNESS

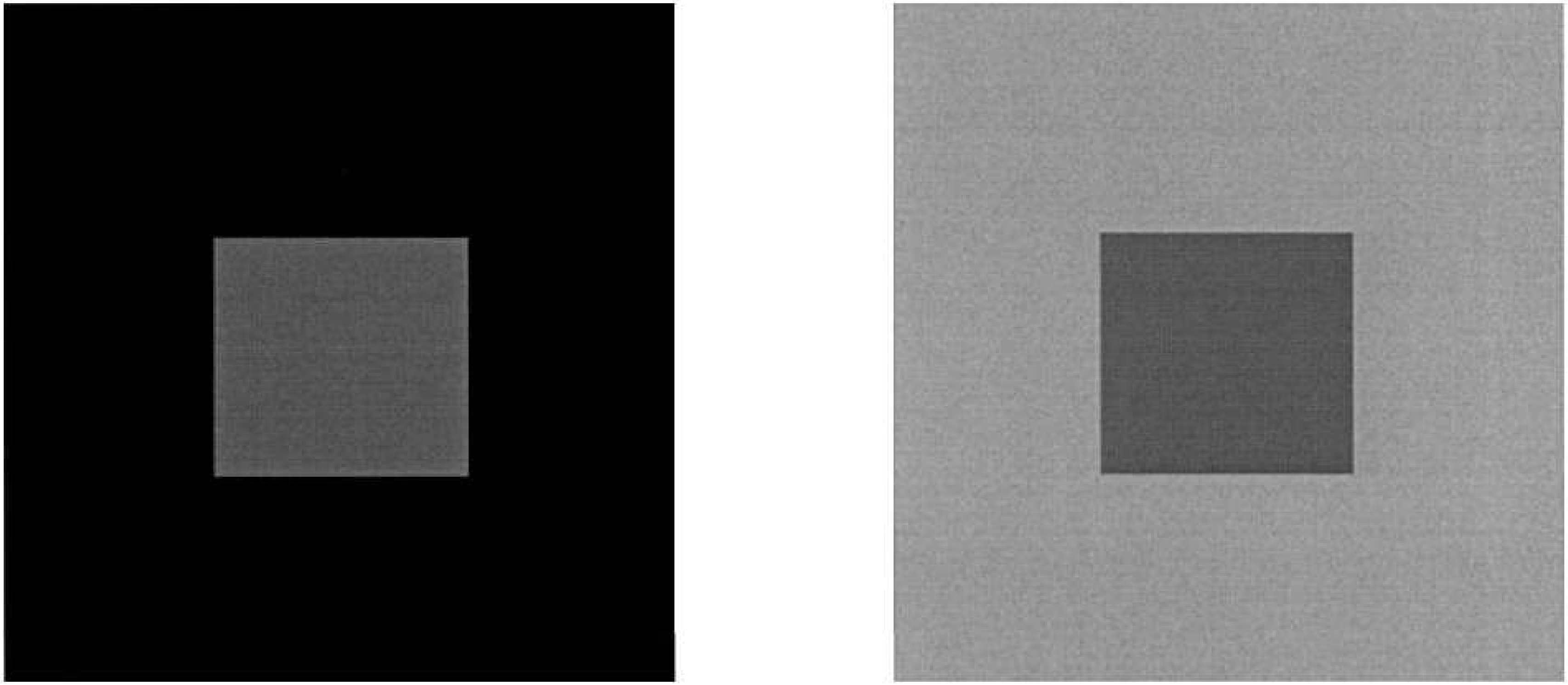

FIGURE 5. REFERENCE-DEPENDENCE IN THE PERCEPTION OF BRIGHTNESS

(The two inner squares are equally bright.)

FIGURE 7. AN ILLUSION OF ATTRIBUTE SUBSTITUTION

FIGURE 7. AN ILLUSION OF ATTRIBUTE SUBSTITUTION

(The two horsies are exactly the same size.)

Note that the illusions persist even after they are pointed out and their mechanisms understood.

Three main elements of prospect theory

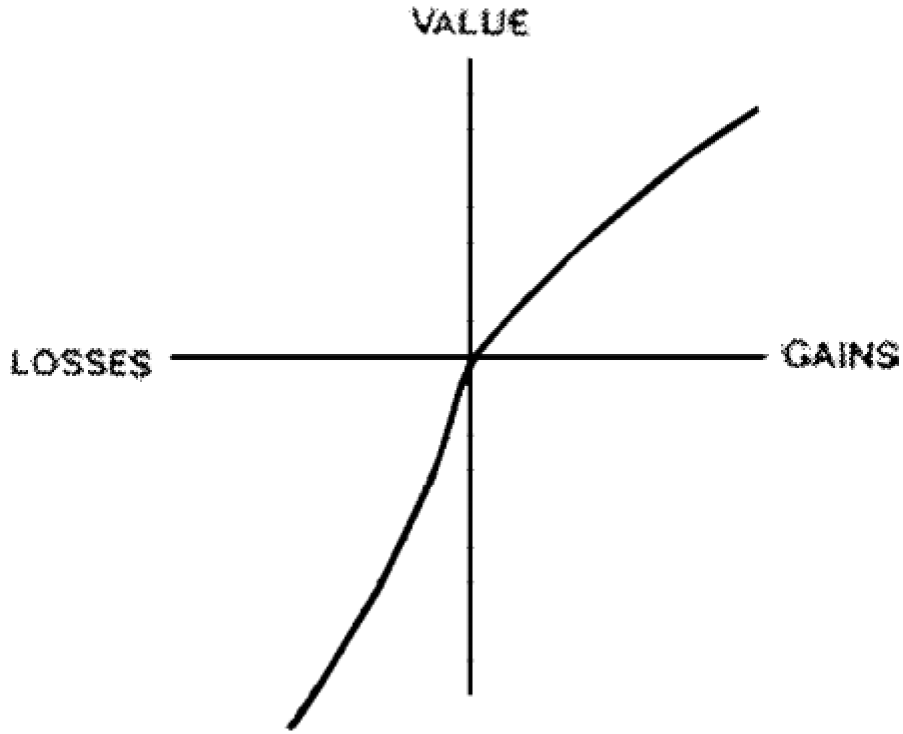

FIGURE 6. A SCHEMATIC VALUE FUNCTION FOR CHANGES

● Loss aversion

Note the kink at 0 (the reference point), so a small decrease below the reference point hurts (in value—KT's word for "utility") 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 around 2 or 3. (Empirically, people seem to be at worst neutral between gains and losses, with a coefficient of 1 as in a neoclassical model: although "gain aversion" is logically just as possible as loss aversion, it is never found.)

● **Diminishing sensitivity**

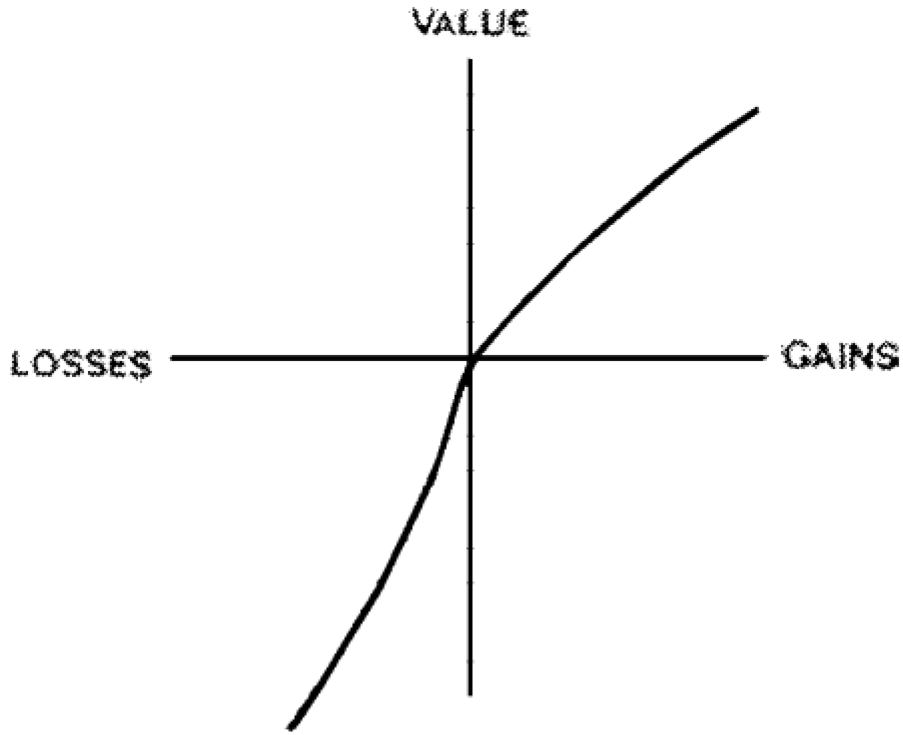

FIGURE 6. A SCHEMATIC VALUE FUNCTION FOR CHANGES

The value function exhibits diminishing marginal sensitivity to losses as well as gains, making it concave for gains but convex for losses.

Because diminishing sensitivity is more relevant to decisions under uncertainty, we'll return to it below.

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

A third feature of prospect theory, nonlinear probability weighting, is a kind of fudge factor by which people are assumed to overweight small probabilities and underweight large ones, so that the value of a risk is $\pi(p)v(x) + \pi(q)v(y)$ rather than $pv(x) + qv(y)$.

Nonlinear probability weighting is less important and less well established empirically than loss aversion and diminishing sensitivity, and will not be discussed here even when we consider decisions under uncertainty.

Return of the mug people

Recall that KKT randomly gave mugs to half the subjects (“owners”) and nothing to the others (“nonowners”).

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

Supply and demand “should” be mirror images of each other. But…

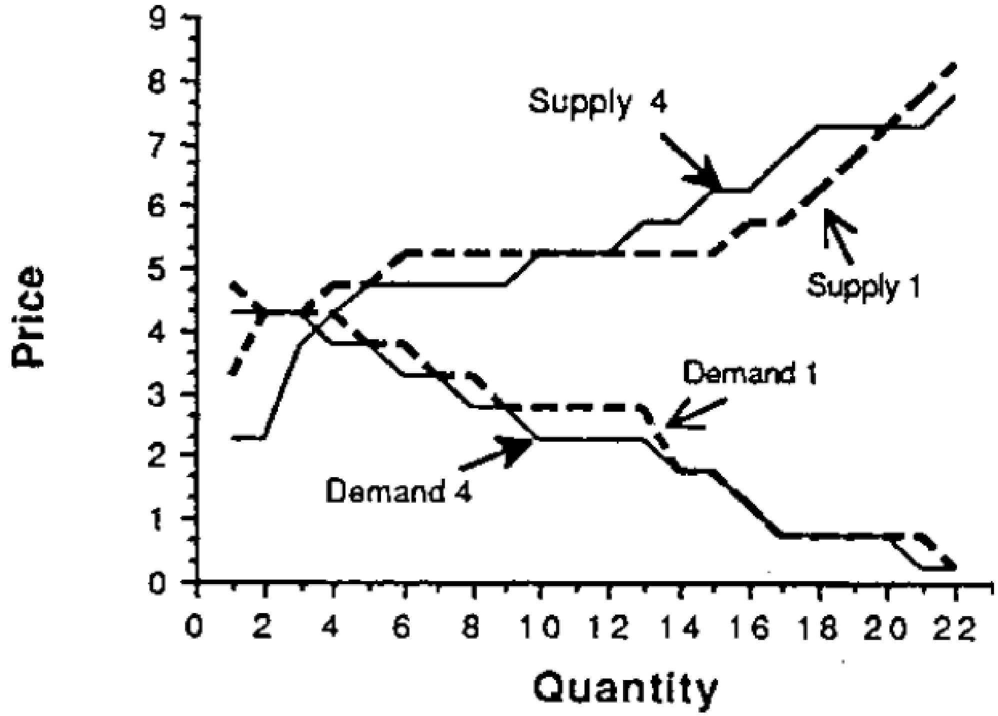

FIG. 1.—Supply and demand curves, markets 1 and 4

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

How do we model this with reference-dependence and loss aversion?

Imagine (unlike Kahneman and Tversky, but like Kőszegi and Rabin (2006 *QJE*)) that people have both ordinary consumption utilities for mugs and money, and gain-loss utilities (which Kahneman and Tversky focused on, as may be appropriate for laboratory experiments with small items).

Assume that subjects have linear consumption utility: value = value of mug (or not) + money, and that owners' and nonowners' consumption utilities for mugs are uniformly distributed between \$0 and \$9 (using the natural money metric).

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 + η × gain-loss utility.

Subjects' reference points are determined by their expectations (as in Kőszegi and Rabin):

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

Nonowners 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 the mug or 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 ≈ KKT's \$7. (A distribution of values makes it easy to generate an entire supply curve as above.)

Demand for mugs

Similarly, a non-owner with mug consumption value $\$v$ who is considering trading $\$m$ of her/his 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 nonowner 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 $m\text{^} = v(1 + \eta)/(1 + 2\eta)$.

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

(A distribution of values makes it easy to generate an entire demand curve as above.)

This explanation links two widespread empirical regularities, the prevalence of gaps (with WTA > WTP) and the prevalence of loss aversion over gain aversion.

The role of expectations

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 nonowner who expects to buy a mug for $\$y$ (so her/his reference point is having a mug but $-\$y$).

These expectations make both sellers and buyers more willing to trade, creating a preference bias, relative to the standard model, in favor of what was expected.

That is the reasoning behind this quotation from Kőszegi and 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 nonowner's decision is just like the one in the "shopping for shoes" example from Koszegi and Rabin's paper, which makes some interesting (though slightly more difficult) further points.

Review of neoclassical rational choice theory continued: Representation of preferences over uncertain outcomes via expected-utility maximization

For choice under uncertainty (viewed as uncertainty over which of a given list of outcomes will happen), if preferences over probability distributions over outcomes satisfy the von Neumann-Morgenstern ("vN-M") axioms:

- Completeness and transitivity, as for the certainty case;

- Continuity in a technically different form ("mixture continuity"); and

- "Independence", which guarantees the separability-across-states of expected-utility maximization

then one can construct a vN-M utility function (cardinal, in a sense), assigning utilities to each possible outcome such that the individual chooses among actions, or equivalently among distributions of outcomes, as if to maximize expected utility.

Example: Suppose that you care only about final lifetime wealth, that your initial wealth w is \$2 million dollars, and you are asked to choose whether to accept a bet (investment opportunity, insurance contract, etc.) that will add either x, y, or z (which could be negative) to your final wealth, with probabilities p, q, or $1 - p - q$ respectively.

Then the vN-M Theorem says that, under the vN-M axioms, 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)

$$pu(w + x) + qu(w + y) + (1 - p - q)u(w + z) > u(w).$$

Just as the utility function used to describe preferences over certain outcomes is a compact way to describe choices in various settings under certainty, the vN-M utility function is a compact way to describe choices in settings involving uncertainty.

As stated, the vN-M Theorem assumes that the probability distributions are objective or at least known.

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

Neoclassical expected-utility maximization and observed choice under uncertainty

The above example bundles two kinds of assumptions:

- That the person's preferences satisfy the vN-M axioms.

- That the person's preferences respond only to levels, in this case of the person's lifetime wealth, as opposed to changes in wealth.

  (I will call this "expected-utility-of-wealth" maximization.)

The vN-M axioms are not completely uncontroversial, and are sometimes systematically violated in observed behavior.

Most notable here (but not covered in lectures; for more detail see http://dss.ucsd.edu/~vcrawfor/142BehavioralUncertaintyLectureSlides08.pdf) are the Ellsberg Paradox, in which people reveal aversion to uncertainty about probabilities ("ambiguity aversion"), violating the assumption that preferences are defined only over probability distributions of outcomes.

(A vN-M person uncertain about probabilities would just calculate the expected probabilities, not otherwise caring about their distributions. This is not explicitly an axiom, but it is built into the framework, which assumes that only the probability distribution of outcomes matters.)

Also notable (and also not covered) is the Allais Paradox, which reveals that people often violate the separability across states assumed in the independence axiom.

Although such violations are important and widely studied, they seem behaviorally less important than violations of the second assumption—that the person's preferences respond to levels rather than changes—which will continue to be our main focus here.

Note first that the assumption that preferences respond only to levels rather than changes is no more logically necessary to use expected-utility maximization to describe choice under uncertainty than it was to use utility-maximization to describe choice under certainty.

It is only a convention, which could be replaced by an alternative if it were found to be useful.

This can be done, as in choice under certainty, simply by expanding or changing the list of things preferences are assumed to respond to.

The vN-M Theorem continues to apply because its axioms are silent on what preferences are about.

To see whether replacing the assumption that preferences respond to levels by allowing them to respond to changes might bring the theory closer to observed behavior, first consider the answers to survey question 1.

(Actually there were two versions of question 1, each answered by half of you: 1a referring to losses and 1b referring to gains.

Comparing answers across versions yields more information than examining them in isolation. Because the versions were assigned randomly and the class is fairly large, differences across versions probably reflect something systematic about people in general, rather than accidental differences across the groups assigned each version.)

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

(either hypothetical as here, or scaled-down but with real payments)

choose to lose $1000 with probability 0.5 rather than $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. Would you choose to receive $500 for sure or to receive $1000 with probability 0.5?

choose to receive $500 for sure rather than $1000 with probability 0.5, suggesting “risk-averse” behavior with respect to gains.

Further (although the survey did not address this), responses to questions like 1a or 1b are approximately independent of bet scale:

People are roughly equally risk-averse for gains, large or small; and roughly equally risk-loving for losses, large or small.

These patterns of risk-loving for losses, risk-aversion for gains, and approximate invariance to the scale of risk in each case are, practically speaking, inconsistent with maximizing the expectation of a utility defined over wealth levels of the usual neoclassical shape.

First, recall that the expected-utility-of-wealth framework captures risk aversion (risk-loving) via concavity (convexity) of the vN-M utility function:

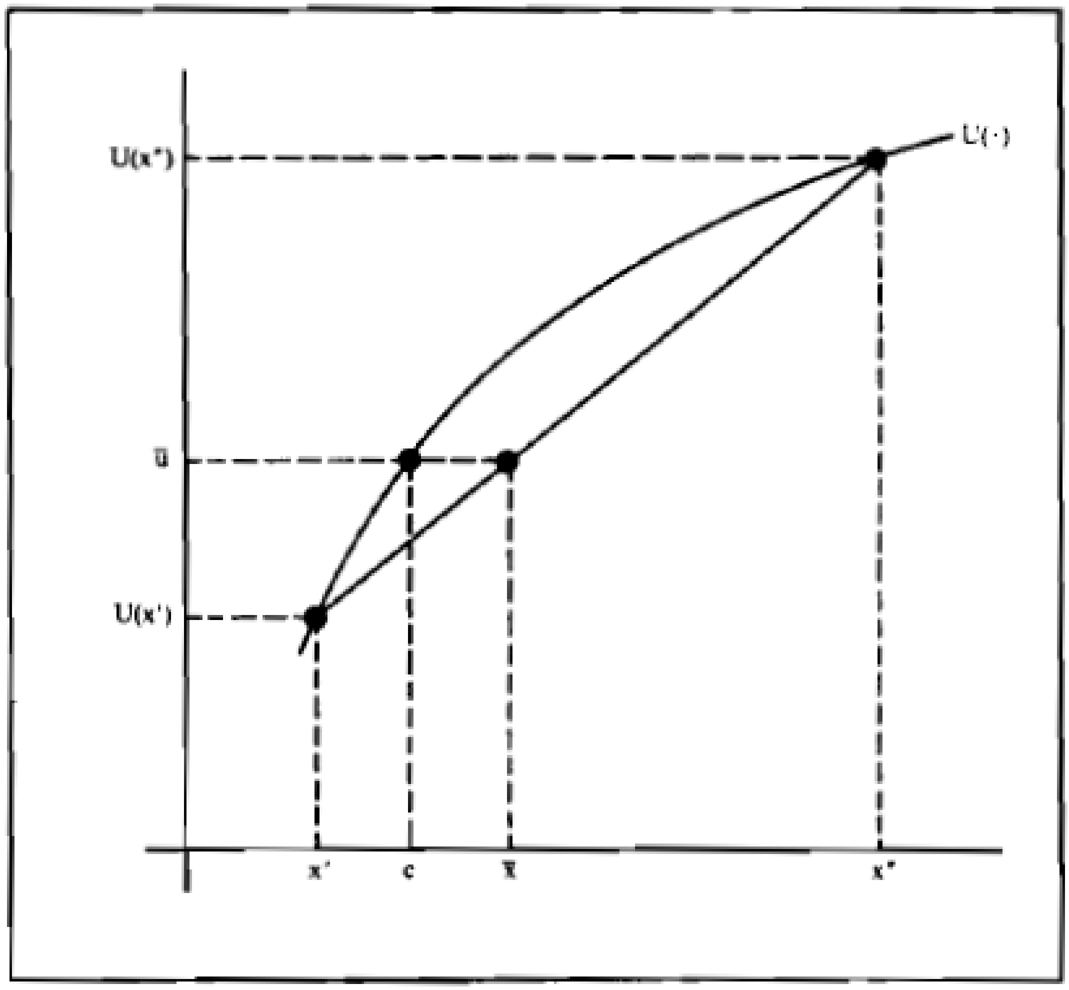

Figure 3 Von Neumann-Morgenstern Utility Function of a Risk Averse Individual

In this framework, if vN-M utility is defined over lifetime wealth levels, and the sums in questions 1a or 1b are small in proportion to lifetime wealth, then choices for 1a and 1b should usually be qualitatively the same—both risk-loving or both risk-averting.

A person with vN-M utility function u(·) and base wealth w will take a 50-50 win σ-lose σ gamble with "risk premium" π if and only if $\frac{1}{2}\,u(w + \sigma + \pi) + \frac{1}{2}\,u(w - \sigma + \pi) \geq u(x)$.

Expanding the left-hand side in a Taylor Series in σ around $\sigma = 0$:

$$u(w + \pi) + \tfrac{1}{2}\,u'(w + \pi) - \tfrac{1}{2}\,u'(w + \pi) + 2(\tfrac{1}{2})2u''(w + \pi)\sigma^2$$
$$= u(w + \pi) + \tfrac{1}{2}u''(w + \pi)\sigma^2 \geq u(w).$$

Expanding the left-hand side in a Taylor Series in π around $\pi = 0$, neglecting the small u''' term, and solving:

$$u(w) + \pi u'(w) + \tfrac{1}{2}u''(w)\sigma^2 \geq u(x),$$
$$\text{which is true if and only if } \pi \geq -\tfrac{1}{2}[u''(w)/u'(w)]\sigma^2.$$

Thus, the risk premium $\pi = -\frac{1}{2}[u''(w)/u'(w)]\sigma^2$ for the bet's utility loss is proportional to $-u''(w)/u'(w)$, the Arrow-Pratt coefficient of absolute risk aversion (normalized to be > 0 if and only if u(·) is concave), and also proportional to σ^2.

("Absolute" risk aversion because $-u''(w)/u'(w)$ measures the utility cost of bets involving absolute changes in w, just as the Arrow-Pratt coefficient of relative risk aversion does for relative changes in w.)

So why are risk-loving for losses, risk-aversion for gains, and approximate invariance to scale of risk "practically speaking" inconsistent with expected-utility-of-wealth maximization?

I stress that they are not *logically* inconsistent with it.

First, assuming continuity and differentiability, the absolute risk aversion function $-u''(w)/u'(w)$ can be any function we want without violating the vN-M axioms or the basic assumption that $u'(w) > 0$.

(Fix a function $-u''(w)/u'(w)$ and integrate to get $u(w)$. The constant of integration doesn't affect expected-utility maximizing choices, so it doesn't matter. Conversely, fixing $u(w)$ implies a unique function $-u''(w)/u'(w)$. The two functions are equivalent representations of preferences.)

Thus we could choose $u(w)$ independently for each person that, given his initial w, is risk-loving for losses and risk-averse for gains, just as in the majority responses for 1a and 1b.

On these grounds, some theorists argue that the patterns noted above are not evidence against the standard model.

But there's no reason why people with widely varying initial w's should all (or most) have $u(w)$'s that just happen to flip from risk-loving to risk-averse at the zero-gain outcome.

We normally assume that preferences are independent of w, not related to it with magical precision.

Aside

It may seem that we could argue that if the people who answered 1a start out as wealthy on average as those who answered 1b, then the gains in 1b tend to make those who answered it richer than the losses in 1a make those who answered it.

If absolute risk aversion is increasing in wealth, people tend to be more risk-averse for gains than for losses, as in the survey responses.

However, the conventional neoclassical assumption (motivated by using the theory to think about observed behavior) is that absolute risk aversion is *de*creasing in wealth.

And even if it were increasing, applications strongly suggest that $-u''(w)/u'(w)$, although it may vary with w, varies far too slowly to differ for moderate (relative to lifetime wealth) gains and losses as radically as they would need to to make people flip from risk-loving for losses to risk-averse for gains.

Die-hard supporters of using the effect of wealth on risk aversion to rescue expected-utility-of-wealth maximization should consider alternative questions posed 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.)

But in terms of probability distributions of final outcomes, these two choices are mathematically identical.

Thus the large flip in the choice distribution 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.

As in the responses to questions 1a and 1b, this appears to make people risk-averse in problem 1 but risk-loving in problem 2.

A similar point is made by survey questions 2a and 2b.

Most people who answer questions like

2a. Would you choose to receive \$3,000 for sure or to receive \$4,000 with probability 0.8?

choose to receive \$3000 for sure.

By contrast, most people who answer questions like

2b. Would you choose to receive \$3,000 with probability 0.25 or \$4,000 with probability 0.2?

choose to receive \$4,000 with probability 0.2.

This by itself is not clear evidence that something other than distributions of final levels matters. But re-frame 2b as a two-stage decision as follows:

In the first stage, with probability 0.75 the process ends with you winning \$0, and with probability 0.25 you move into the second stage.

In the second stage, you choose between receiving \$4,000 with probability 0.8 and \$3,000 for sure. (Your choice here must be made before the outcome of the first stage is known.)

This is mathematically identical to the original choice 2b, but here, unlike in 2b, most people choose \$3,000 for sure in the second stage.

Once they see the chance of getting \$3,000 for sure, they think about the risk differently: Although the inference is now complicated by intertemporal issues, the example suggests that more than the distribution of final levels matters.

End of aside

Scaling: Using expected-utility-of-wealth maximization to explain choices involving large as well as small risks

More subtle evidence against expected-utility-of-wealth maximization is implicit in the frequent observation that people seem much more averse to small risks than such maximization would predict, given their willingness to take larger risks.

For example, we saw above that with expected-utility-of-wealth maximization, the risk premium $\pi = -½[u''(w)/u'(w)]\sigma^2$ grows not with the scale of the bet σ but with its variance σ^2.

This implies that risk-averse people are approximately risk-neutral for small bets—“second-order risk aversion”—but disproportionately risk-averse for larger bets.

With a differentiable, increasing vN-M utility function over wealth, a risk-averse person may turn down some more than fair bets, because the “cost” of a large risk may outweigh its positive expected return.

But if such a 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.

As a result, people turning down small bets who have globally risk averse vN-M utility functions over final wealth most be insanely risk-averse over large more-than-fair bets.

But most people’s behavior with respect to large and small risks suggests that they have “first-order” risk aversion, with risk premia approximately proportional to scale.

Rabin and Thaler (2001 *Journal of Economic Perspectives*) 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.

(Rabin again:)

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. [Me: Here he means radical changes in absolute risk aversion.]

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 $5,000 in pre-tax income per year, say—then he would value income only at most ≈ 0.8% ($\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: Compounding small risks

People also seem to be more comforted by compounding small risks than expected-utility-of-wealth maximization predicts.

Rabin-Thaler (*JEP* 2001):

> 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.

End of aside

Reference-dependent preferences

Reference-dependent preferences again suggest a natural, plausible, and parsimonious explanation of these phenomena.

As before, we replace the assumption that the outcomes over which utility is are defined are limited to lifetime consumption or wealth bundles, expanding the space over which preferences are defined to include a reference point as well.

Preferences are then defined not over levels, but over changes in the sense of gains or losses relative to the reference point.

It again seems natural to take the reference point as the status quo before the choice (but the alternative of taking it to be determined by expectations is also considered).

It also seems natural to assume that subjects consider the choice situation in isolation, without trying to integrate it into a lifetime consumption plan ("narrow bracketing").

Aside:

The most plausible escape routes other than reference-dependence are closed off:

● 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 (2009 *Econometrica*) show that the reactions are separate from those that underlie the Allais paradox.

● And kinks can’t be ubiquitous enough to save expected-utility-of-wealth maximization because a concave vN-M utility function must be differentiable almost everywhere.

End of aside

In prospect theory, a person's attitudes toward risk are jointly determined by her/his degree of loss aversion and/or diminishing sensitivity.

If we agree to ignore nonlinear probability weighting, we can re-run the vN-M Theorem to justify expected (prospect theory) value maximization as a representation of reference-dependent preferences over uncertain outcomes.

The key point is that the Theorem works for preferences defined over anything, including changes relative to a reference point rather than levels.

With a tractable model of the reference point and a reasonable specification of diminishing sensitivity (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.

Although the behavioural literature sometimes makes a big deal about diminishing sensitivity and even nonlinear probability weighting, and they are realistic and important for some applications, most of the action in prospect theory comes from reference-dependence and loss aversion.

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.

Further, it's possible to make sense of many phenomena using a piecewise linear value function with a coefficient of loss aversion of approximately 2.

(The evidence suggests that while some people are not loss-averse, in which case prospect theory is close to standard expected-utility theory, *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.)

Loss aversion and "first-order" risk aversion

Unlike expected-utility preferences, prospect theory preferences with loss aversion have a built-in, portable kink at the reference point which allows the theory easily to accommodate first-order risk-aversion, even with an otherwise differentiable value function.

Ignore diminishing sensitivity for simplicity, so that the value function $v(\sigma)$ is piecewise linear: linear except for a kink at the reference point.

Normalize the reference point to 0, with $v(0) = 0$, and set the coefficient of loss aversion at 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 "first-order" risk aversion because the risk premium $\pi = \sigma/2$ grows linearly with the scale of the bet σ.

Recall that, by contrast, expected-utility preferences made the risk premium grow not with the scale of the bet σ but with its variance σ^2.

(If we allowed diminishing sensitivity and 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.)

Diminishing sensitivity

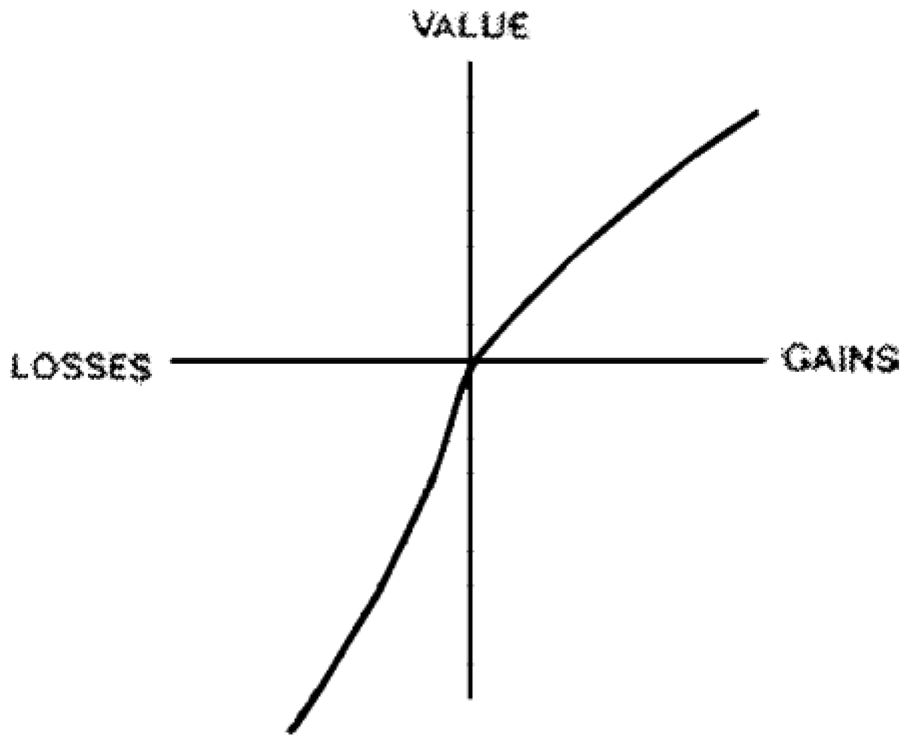

FIGURE 6. A SCHEMATIC VALUE FUNCTION FOR CHANGES

Diminishing sensitivity adds nuance to the person's risk preferences, with the value function concave for gains but convex for losses, so s/he is more risk-averse for gains, other things equal.

Risk aversion is still dominated by the first-order effects of loss aversion for decisions that imply a positive probability of crossing the reference point, and this can easily outweigh the second-order risk-loving behavior associated with convexity of the value function for losses.

Kahneman and Tversky (1979 *Econometrica*) argue that diminishing sensitivity reflects a 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....

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 that most people who answer questions like

1a. Would you choose to lose $500 for sure or to lose $1000 with probability 0.5?
choose to lose $1000 with probability 0.5 rather than $500 for sure, suggesting "risk-loving" behavior with respect to losses.

And that by contrast, most people who answer questions like

1b. Would you choose to receive $500 for sure or to receive $1000 with probability 0.5?
choose to receive $500 for sure rather than $1000 with probability 0.5, suggesting "risk-averse" behavior with respect to gains.

Sample applications

Thaler and Johnson (1990 *Management Science*) propose a prospect theory explanation of the phenomenon that race-track bettors tend to bet more on long shots near the end of the day.

Here it’s natural to take the period over which gains and losses are evaluated (the “bracket”) as the day at the track, and to take the reference point as breaking even.

Loss aversion without diminishing sensitivity—a piecewise linear value function—makes people who have little chance of crossing the reference point with the next bet (whether because they’re way ahead or way behind) less risk-averse than people who do have a significant chance.

This generates both the “house-money” effect, in which people are more willing to bet when they’re ahead: and the “break-even” effect, in which people are more willing when they’re behind.

The non-monotonicity is a challenge to expected-utility explanations, particularly as the explanation must work for a wide range of initial wealth levels.

Adding diminishing sensitivity to loss aversion makes people who are behind somewhat more willing to bet than people who are ahead. (Jaimie Lien (2009) analyzes a new dataset on casino betting from this point of view, finding both loss aversion and diminishing sensitivity.)

The break-even effect seems to make long shots more attractive to most bettors (most of whom are losers) near the end of the day.

This effect is strong enough to make betting on the favorite to show in the last race profitable.

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.

In a standard analysis, sellers’ optimal asking prices should be approximately independent of what they paid for the house, other things equal: It’s a sunk cost, new buyers’ values are independent of what the seller paid, and seller wealth effects are not large enough to plausibly explain the large differences observed.

But in the data there are dramatic differences:

Sellers who are selling their condos for a nominal loss relative to their buying price charge a higher price than those selling without a loss—on average by 35% of the average difference between the optimal price and the price at which they bought it.

Investor sellers exhibit less of a difference than owner-occupier sellers, but still have some.

Genesove and Mayer carefully rule out other possible explanations, leaving loss aversion.

Here 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 against gains from selling your Rolls).

It’s natural to take the reference point as breaking even relative to what you paid for the house (apparently without controlling for inflation).

Cicchetti and Dubin (1994 *Journal of Political Economy*) study people's decisions to buy insurance against damage to their home telephone wiring: people in their sample paid almost twice the expected cost to insure against losses typically less than $100.

Justin Sydnor, "Abundant Aversion to Moderate Risk: Evidence from Homeowners Insurance," 2006, http://wsomfaculty.case.edu/sydnor/deductibles.pdf , studied people's choices of deductibles for home insurance.

His customers chose between four deductibles: $100, $250, $500, and $1,000. (His data also include house characteristics, premiums, claims, which reveals how much they would have paid and/or received had they chosen a different deductible.)

Almost no one chose the $100 deductible, but the other levels were often chosen. People overpaid for lower deductibles by a factor of 5.

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?

You say ferrets and tea kettles are trivial examples?

Okay, how about stock or labor markets? Camerer, Babcock, Loewenstein, and Thaler (1997 *QJE*) 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.

Odean (1998 *Journal of Finance*) and Meng (2009) use prospect theory to explain a strong and widely observed "disposition effect", the tendency for stock market traders to sell winners more readily than losers, other things (including estimated expected future returns) equal.

Camerer et al. (1997 *QJE*) also study the labor supply of New York City cabdrivers, who 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:

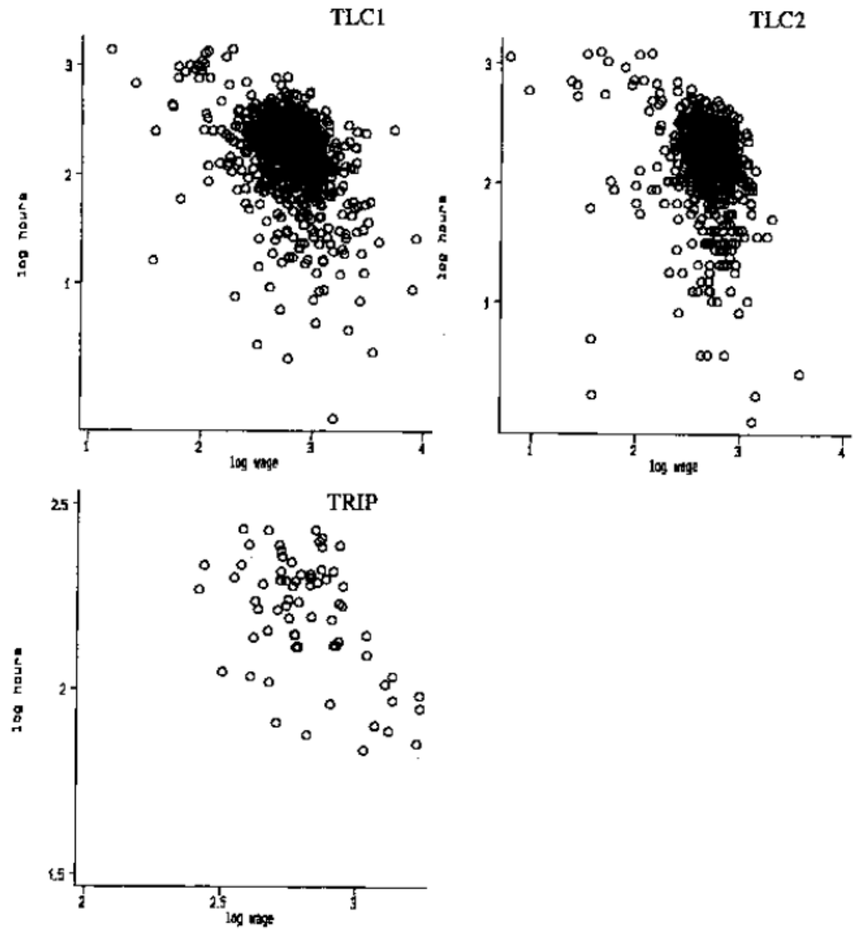

FIGURE I
Hours-Wage Relationships

The elasticities look 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, relatively easily, by working longer that day.

Camerer et al. (1997 *QJE*; see also Kőszegi and Rabin 2006 *QJE*) 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 (Kőszegi 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.

But Farber (2005 *Journal of Political Economy* and 2008 *AER*) challenged this explanation.

Crawford and Meng (2009) challenge the challenge, using Kőszegi and Rabin's (2006 *QJE*) model to explain the tendency for drivers' daily hours and earnings to bunch around expectations-based targets, and applying the model econometrically to Farber's data and showing that it does yield a credible account of Farber's cabdrivers' daily labor supply decisions.

Unresolved issues: “mental accounting” and “narrow bracketing”

Mental accounting and narrow bracketing are two important issues that have been 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, or 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 (but see Kőszegi and Rabin 2009 *AER*). The questions are empirical but fortunately there are regularities in the data to guide assumptions about them.