Unexpected Utility: Experimental Tests of Five Key Questions about Preferences over Risk*

James Andreoni
University of California, San Diego and NBER

William T. Harbaugh
University of Oregon

December 2009
(Current version April 3, 2010)

Abstract

Experimental work on preferences over risk has typically considered choices over a small number of discrete options, some of which involve no risk. Such experiments often demonstrate contradictions of standard expected utility theory. We reconsider this literature with a new preference elicitation device that allows a continuous choice space over only risky options. Our analysis assumes only that preferences depend on the probability p and prize x, $U = u(p, x)$. We then allow subjects to choose p and x continuously on a linear budget constraint, $r_1 p + r_2 x = m$, so that all prospects with a nonzero expected value are risky. We test five of the most importantly debated questions about risk preferences: rationality, prospect theory asymmetry, the independence axiom, probability weighting, and constant relative risk aversion. Overall, we find that the expected utility model does unexpectedly well.

*We would like to thank Michele Cohen, Shachar Kariv, Justin Rao, Charles Sprenger, and Jean-Chrisotphe Vargnaud for helpful comments, and Daniel Gelman and Marion Price expert computer programming and research assistance. This work was supported by NSF grants SES-0551296 (Andreoni) and SES-0112157 (Harbaugh).

1 Introduction

Hundreds of theoretical and experimental papers have been written about preferences over risk. Of the experimental papers, the vast majority study choices over a small, finite sets of lotteries, one of which is often a sure thing. These papers frequently result in contradictions of the standard model of expected utility. The most famous of these is the Allais Paradox (1953). Allais argues (2008, page 4) that "in the neighborhood of certainty" the independence axiom breaks down and individuals show a disproportionate "preference for security," while away from certainty the standard model of expected utility should hold. This has become known as the certainty effect, and is well documented in the literature.

Of equal prominence is prospect theory, both the asymmetry around gains and losses of Kahneman and Tversky (1979), known as loss-aversion, and cumulative prospect theory summarized by probability weighting (Quiggin 1982, Tversky and Kahneman, 1992, Tversky and Fox, 1995, Prelec, 1998). Measures of loss aversion often come from comparing choices over discrete sets of gambles that look quite different. In life, gambles are commonly continuous (how much savings to put in stocks versus bonds, how fast to drive, how long to continue search). This leads one to wonder whether presentation complicates the choice task and invites the use of simplifying heuristics.[1] In addition, models of probability weighting are built from data that elicits an individual's certainty equivalent, that is, the certain amount that would make a subject indifferent to a particular lottery. In light of the certainty effect of Allais, however, using certainty equivalents builds the certainty effect into the data by mixing a certain outcome with risky outcomes, and thus may generate misleading conclusions about choices over only risk.[2]

We develop a new elicitation method and reconsider the five key questions that have dominated the literature on risk preferences. Rather than starting with a particular anomaly or theory, we assume simply that individuals have some preferences over winning (or losing) an amount x with probability p, so $U = U(p, x)$. When x is a gain, both p and x are "goods," and most people would prefer higher values for both. When $x < 0$ is a loss, then p and $|x|$ are "bads," and people prefer less of each. If we were to study preferences for goods and bads other than risk, we would begin by optimizing a utility function, $U(p, x)$, subject to a linear budget constraint, say $r_1 p + r_2 x = m$. By changing "prices" (r_1, r_2) and "income" m, we can learn about the consistency of preferences and the shape of the utility function.[3] This is where we begin with our study.

Our approach will allow us to learn about preferences in a new light. Choices are on a continuum, are simple, and always involve risk. More importantly, we provide strict tests

[1] This has been argued by Leland (1998,2009) and Rubinstein (1988), for instance.

[2] See Andreoni and Sprenger (2009b) for a demonstration of the possible bias from using certainty equivalents to measure constant relative risk averse utility.

[3] This study is distinct from the important prior work of Choi, Fisman, Gale and Kariv (2007). They fix probabilities at, for example, 50-50 chances of winning prize x_1 or x_2, and allow subjects to allocate the prizes on a linear budget, $q_1 x_1 + q_2 x_2 = m$. Their "graphical interface" program generates fifty random pairs of prices (q_1, q_2) per subject. Their approach also allows for tests of consistency of preferences. Our study allows people to make trade-offs between higher prizes at lower probabilities, and thus can identify theories that make assumptions about these trade-offs, including the independence axiom, and probability weighting. Moreover, we consider losses as well as gains, so can address prospect theory asymmetry and loss aversion.

that can inform us on five most important questions on risk, and can falsify expected utility and its alternatives. First, does a quasi-concave utility function exist that could rationalize the data? Second is prospect theory asymmetry: is there loss aversion around a reference point of zero, and are preferences risk averse on gains and risk loving on losses? Third is the independence axiom (von Neuman and Morgenstern, 1944), which is the workhorse of standard expected utility theory, satisfied? Fourth is probability weighting, the chief alternative to the independence axiom, supported? Finally, is an assumption of constant relative risk averse utility too restrictive to be considered a sensible simplification?

We titled our paper "unexpected utility" for three reasons. First, we began this project without expectations, but only with a hope of getting refutable statements of core assumptions, and more precise estimation of preferences. Second, we do not need to assume parametric forms of expected utility to test some of its tenets–our maintained hypothesis is simply utility, not expected utility. This allows a very powerful test of the theory. And third, we think many people will find our results unexpectedly supportive of neoclassical expected utility theory.

Our answers to the five questions are, first, that choices over gains are largely consistent with a model of rational choice where people treat the probability and the prize as goods. Decisions on losses, by contrast, are much more noisy, perhaps because the decision task over losses is less familiar. Second, nonparametric tests show a significant minority of subjects satisfy the assumptions of prospect theory asymmetry—risk aversion on gains and risk loving on losses. Performing standard parametric analyses on individuals, however, causes the relationship to weaken, while aggregate analysis retains only loss aversion and not risk loving on losses. Third, the data are strikingly consistent with the independence axiom, which is our strongest support for the neoclassical model. By contrast, our data reject probability weighting, which we speculate is due to the fact that the probability weighting model is typically fit using data that conflate preferences over risk with the certainty effect. Finally, we show that CRRA preferences restrict $U(p, x)$ to be a Cobb-Douglas utility function. We find that our data fit this restriction with remarkable precision. Overall, our method of offering choices over a continuum of lotteries, all of which involve risk, yields results that are unexpectedly consistent with expected utility.

The next section explains how our methodology generates tests of the five questions. Section 3 describes the experiment we implement. Section 4 reports the aggregate data, while sections 5 to 9 present the tests of the five assumptions. Section 10 is a discussion and conclusion.

2 Preferences for Risk: The Five Questions

In this section we return to the fundamentals of choice under uncertainty. At each step we add more structure, and provide a test that could reject the restrictions.

2.1 Question 1: Are Preferences over Risk Rational?

Consider a lottery that has a probability p of winning $x > 0$, and wins zero otherwise. Then the expected value of the gamble is

$$EV = px$$

Individuals do not generally maximize EV unless they are indifferent to risk. Instead, assume that people have preferences over combinations of p and x,

$$U = U(p, x) \tag{1}$$

where both p and x are "goods" and $U(p, x)$ is continuous and quasi-concave. Suppose we offer p and x on a linear budget, say $r_1 p + r_2 x = m$, where r_1, r_2, and $m > 0$. That is, to get a bigger prize, one has to accept a smaller chance of winning it. Then allow a person the to optimize (1):

$$\max_{p,x} U(p, x) \tag{2}$$
$$\text{s.t. } r_1 p + r_2 x = m$$
$$0 \leq p \leq 1, \ x \geq 0.$$

We now have our first and weakest restriction: If preferences are rational and well-behaved, then (p, x) choices should satisfy the axioms of revealed preference.

If rationality holds, this framework allows nonparametric identification of risk neutrality, risk aversion, and risk loving behavior, even without invoking the independence axiom. A risk neutral person would maximize expected value, $U(p, x) = px$, selecting, say, (p_N, x_N). A risk averse person, by contrast, would prefer a bundle (p_A, x_A) that, relative to (p_N, x_N), has a higher chance of a lower prize: $p_A > p_N$ and $x_A < x_N$. Conversely, a risk lover prefers a lower chance of a higher prize, that is (p_L, x_L) where $p_L < p_N$ and $x_L > x_N$. This is illustrated in Figure 1.

Suppose instead that $x < 0$ is a loss. Now p and $|x|$ are both "bads." This means a person will most prefer corner solutions. Choices over most preferred points, therefore, will not reveal much about preferences. If, however, we were to ask people to *minimize* utility, then analysis like that above applies, except that it is inverted. In particular,

$$\min_{p,x} U(p, x)$$
$$\text{s.t. } r_1 p + r_2 |x| = m$$
$$0 \leq p \leq 1, \ x \leq 0$$

will tend to have an interior solutions as long as $U(p, x)$ is sufficiently convex. Choices will adhere to axioms of choice, but rather than revealing a preference they will be revealing an *aversiveness*. Moreover, if (p_N^ℓ, x_N^ℓ) characterizes the expected value minimizing gamble along this budget constraint, then a risk averse utility minimizer will choose $p_A^\ell > p_N^\ell$ and $x_A^\ell < x_N^\ell$, while a risk loving expected utility minimizer would choose $p_L^\ell < p_N^\ell$ and $x_L^\ell > x_N^\ell$. These are summarized in Figure 2.

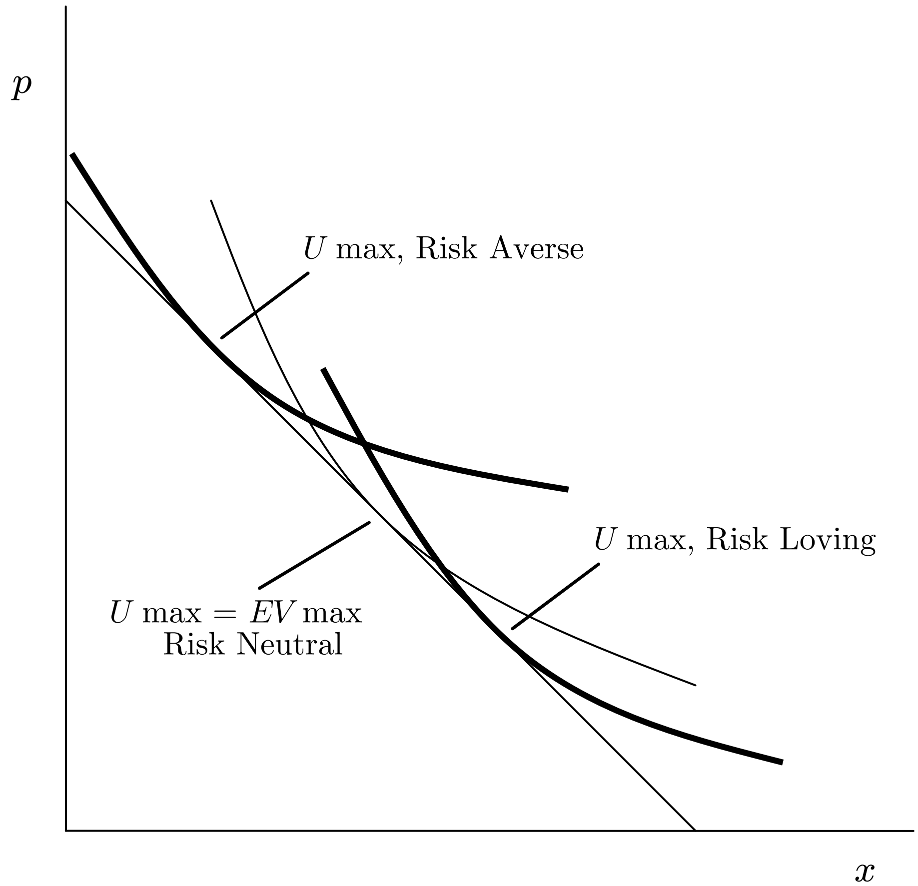


Figure 1: Maximizing $U(p, x)$ subject to a linear budget $r_1 p + r_2 x = m$, and $x > 0$. Comparing risk averse and risk loving preferences to expected value maximization.

2.2 Question 2: Are Preferences Risk Averse on Gains and Risk Loving on Losses?

In an extremely influential body of work, Kahneman and Tversky showed convincingly that utility is better measured relative to some reference point. In risk, that reference point is assumed to be no gain or loss. Somewhat more controversial is their conclusion that, while utility over gains is risk averse, utility over losses is risk loving.[4] The graph of utility over changes in consumption looks S-shaped, with a kink at zero change.

Our methodology will allow us to identify both non-parametrically and parametrically whether individuals exhibit this asymmetry.

[4] For example, see Harbaugh, Krause and Vesterlund (Forthcoming).

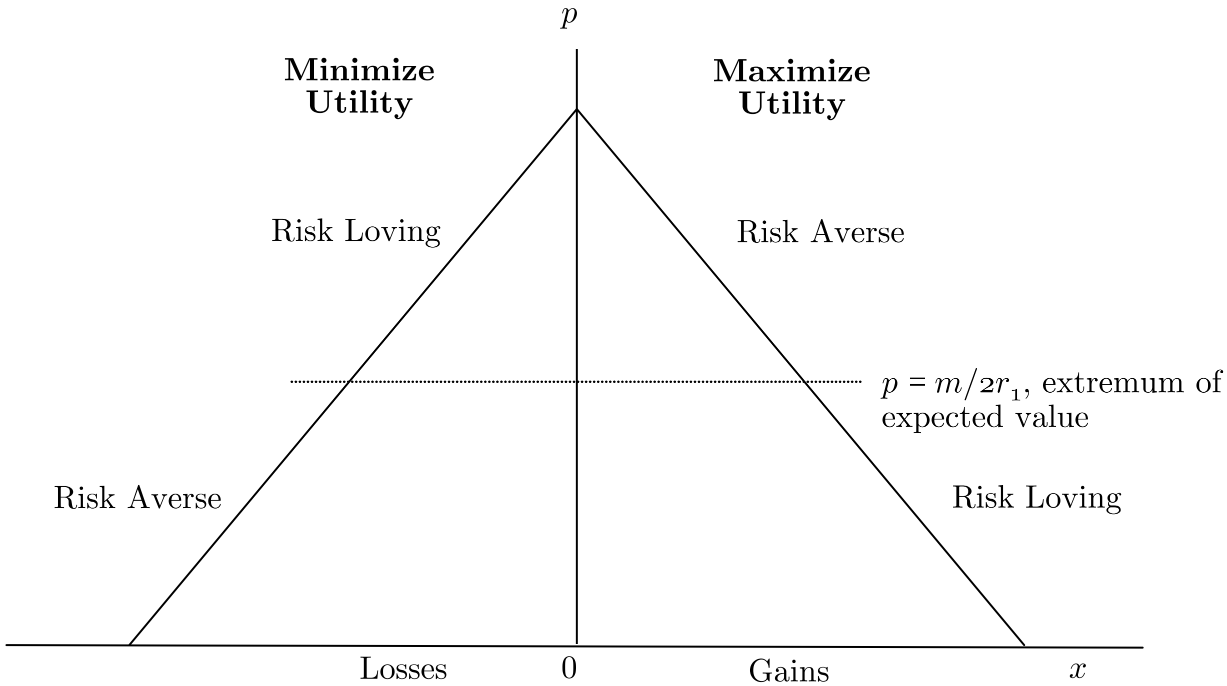


Figure 2: Choices that reveal risk aversion and risk loving behavior for both gains and losses, without relying on the expected utility framework.

2.3 Question 3: Does the Independence Axiom Hold?

A general utility function like (1) is of limited value without additional structure. The assumption at the center of expected utility theory is the independence axiom.[5] This axiom implies that the utility an individual experiences from consuming x is independent of the probability of consuming x. It allows us to write utility as linear in p,

$$U(p, x) = EU = pu(x) \tag{3}$$

with $u(0) = 0$ by assumption.

Solving the optimization problem with expected utility has a powerful and easily testable prediction. Consider marginal conditions of optimizing (3) subject to $r_1 p + r_2 x = m$:

$$\frac{pu'(x)}{u(x)} = \frac{r_2}{r_1}. \tag{4}$$

Cross multiply this to get $r_1 pu'(x) = r_2 u(x)$. From the budget constraint, substitute out $r_1 p$:

$$(m - r_2 x)u'(x) = r_2 u(x).$$

[5]Formally, the independence axiom states that if an agent is indifferent between simple lotteries $L1$ and $L2$, the agent is also indifferent between $L1$ mixed with an arbitrary simple lottery $L3$ with probability p and $L2$ mixed with $L3$ with the same probability p.

This shows that the solution for x is only a function of r_2 and m, $x = x(r_2, m)$, and not a function of r_1. That is, sketching offer curves as the price of p changes, while r_2 and m are constant, should yield vertical lines. This provides a strict test of the independence axiom.

2.4 Question 4: Does Probability Weighting Hold?

The leading alternative to the independence axiom is probability weighting. The hypothesis is that individuals behave as if the probability were really transformed by a weighting function $w(p)$, such that $w(0) = 0$, $w(1) = 1$, but $w(p) > p$ for p close 0, and $w(p) < p$ for p close to 1 with $w(p) = p$ at some intermediate value. If $w(p)$ is differentiable, this means that $w' > 0$ for all p, $w' > 1$ for both p near 0 and near 1. An example of a typical probability weighting function is shown in Figure 3.

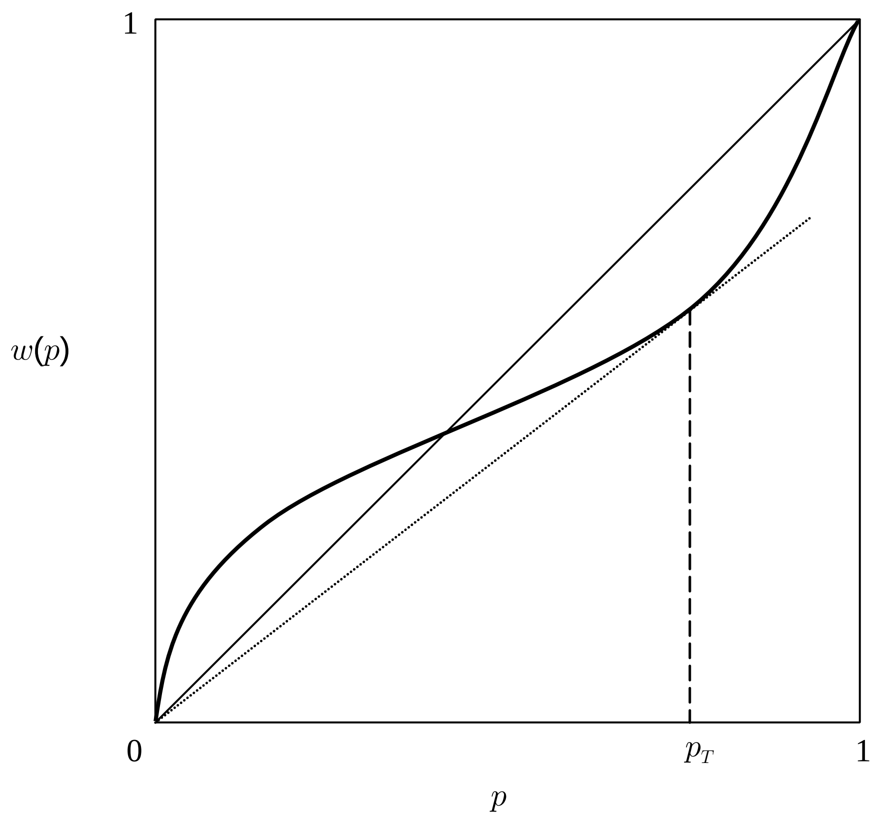


Figure 3: A Typical Probability Weighting function $w(p)$.

The $w(p)$ function was originally derived by Tversky and Fox (1995) by eliciting certainty equivalents from individuals, assuming a CRRA utility of $u = x^{0.88}$, and then solving for the value of $w(p)$ that can justify the certainty equivalents. Given the power of the Allais certainty effect, it may be problematic to infer preference for risk when one of the options being compared (the "equivalent" certain amount) is not risky.[6] It seems reasonable, therefore, to

[6]See Andreoni and Sprenger (2010*c*) for detailed arguments on this point, and for a review of the back-

ask if $w(p)$ takes the same shape when all prospects involve risk.

In our framework, probability weighting means individuals solve

$$\max_{p,x} w(p)u(x)$$
$$\text{s.t.} \quad r_1 p + r_2 x = m,$$

yielding marginal conditions

$$\frac{w(p)u'(x)}{w'(p)u(x)} = \frac{r_2}{r_1}.$$

What falsifiable predictions follow from this solution? Assume that the weighting function is smooth and differentiable, as shown in Figure 3.[7] Of particular importance will be the p labeled p_T in Figure 3. At p_T a line from the origin is tangent to the lower edge of the weighting curve. For $p < p_T$, we know that $w(p) > w'(p)p$. Thus $w(p)/w'(p) > p$. Let (p^w, x^w, r_1^w) describe a choice on an offer curve and relevant price, assuming probability weighting, for any $p < p_T$. Then

$$\frac{p^w u'(x^w)}{u(x^w)} < \frac{w(p^w)u'(x^w)}{w'(p^w)u(x^w)} = \frac{r_2}{r_1^w}.$$

Rearrange and substitute in the budget to get

$$\frac{(m - r_2 x^w)u'(x^w)}{u(x^w)} < r_2 \tag{5}$$

Next let (p^I, x^I) be the solution at this price assuming the Independence Axiom. Then

$$\frac{(m - r_2 x^I)u'(x^I)}{u(x^I)} = r_2. \tag{6}$$

Let $\phi(x) = (m - r_2 x)u'(x)/u(x)$. It is trivial to show that $\phi'(x) < 0$, as long as $u'' < 0$. Combine (5) and (6) to find that $\phi(x^w) < \phi(x^I)$, which means $x^w > x^I$ for all $p < p_T$.

Now consider those $p > p_T$. In this case $w(p) < w'(p)p$. Thus $w(p)/w'(p) < p$. Again let (p^w, x^w, r_1^w) describe a choice on an offer curve and relevant price, assuming probability weighting, for $p > p_T$. Then

$$\frac{p^w u'(x^w)}{u(x^w)} > \frac{w(p^w)u'(x^w)}{w'(p^w)u(x^w)} = \frac{r_2}{r_1^w}.$$

Applying the same logic as above, we get that $\phi(x^w) > \phi(x^I)$ and so $x^w < x^I$ for $p > p_T$. Similarly, $x^w = x^I$ for $p = p_T$.

As shown under question 3, the independence axiom implies that x^I is the same for all prices r_1. Let r_{1T} be the price at which the offer curves under probability weighting and the independence axiom cross, that is, both demand $p = p_T$. Then these results indicate that

ground literature.

[7] The same results follow under more general assumptions. For instance, if it is assumed that $w(p)$ is piecewise linear, but maintains the general property of having a slope greater than 1 at both ends, the same results easily follow.

for prices $r_1 < r_{1T}$, that is $p > p_T$, the offer curve under probability weighting is to the left of that under the independence axiom. For prices $r_1 > r_{1T}$, that is $p < p_T$, the offer curve under probability weighting is to the right of that under the independence axiom.

This gives us a clear test of an alternative to the independence axiom. If the probability weighting holds as in Figure 3, then offer curves should slope back, as in Figure 4.

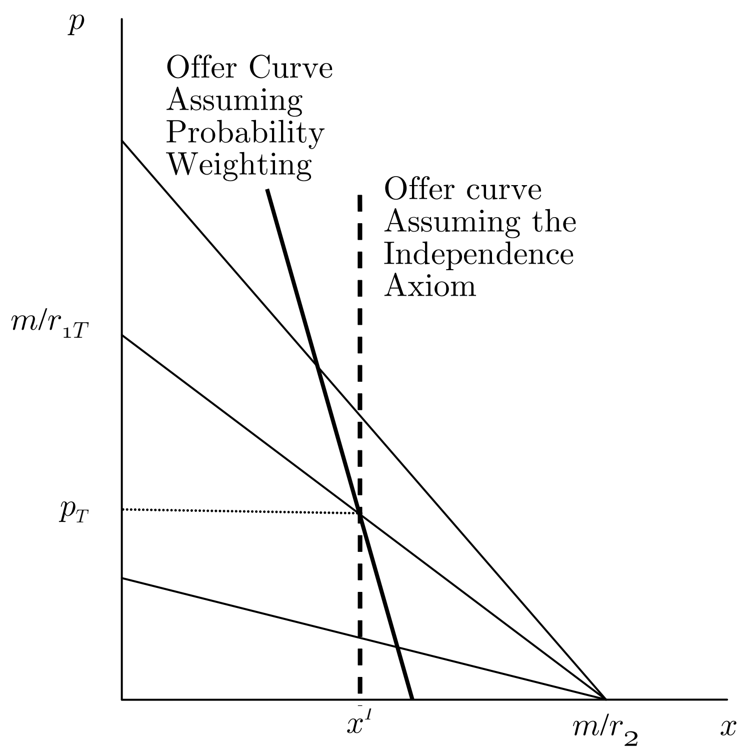


Figure 4: Offer curves assuming individuals satisfy the Independence Axiom, or adhere to Probability Weighting

2.5 Question 5: Is CRRA Utility Too Restrictive?

Assuming a functional form for utility is often necessary for analysis. Clearly the most popular functional form is CRRA:

$$u(x) = x^{\alpha}. \tag{7}$$

CRRA utility has the clear advantage of being a single-parameter function, so can be measured off of a single observation, as is done in Holt and Laury (2002).[8] In our framework, assuming CRRA utility is the same as assuming that expected utility is Cobb-Douglas:

[8] See also Glenn W. Harrison, Morten I. Lau, and E. Elisabet Rutström (2007). The Holt-Laury method observes a single crossing point in a series of binary choices that incrementaly change the risk. The crossing point can be seen as the single observation that identifies a narrow interval of values for α. Harrison, et al. improve on this by iterating the method in order to narrow the range of estimates further.

$U(p,x) = px^{\alpha}$. Maximizing this subject to $r_1 p + r_2 x = m$ yields demands of the familiar Cobb-Douglas utility form:

$$p(r_1, m) = \frac{1}{1+\alpha}\frac{m}{r_1}, \qquad x(r_2, m) = \frac{\alpha}{1+\alpha}\frac{m}{r_2}.$$

As with all Cobb-Douglas demands, the budget shares for each demand ($r_1 p/m$ and $r_2 x/m$) are constant, and the elasticity of substitution between p and x is constant and equal to minus one. We can easily test these restrictions with our data.

3 Experimental Protocol

The experiment was conducted in the Economics Laboratory at UCSD in the fall of 2008. Subjects participated in groups of 20 to 24, with a total of 88 subjects. The experiment was presented by computer. Each session lasted about 75 minutes and subjects earned an average of $22.10 (s.d. 3.91).

Subjects were each given $20 as a show up fee, and were told that they could add to this or lose some or all of it during the study.[9] Each subject made choices over 14 budget sets for gains and 14 for losses. We ran four sessions of the experiment, in two of which subjects made choices over all the gain budget sets first, and in two of which they made choices over all the loss sets first.[10] Within the loss and gains groups, budgets sets were presented in a random order to each subject. Before making any decisions subjects were told that one of the budget sets would be selected at random and the gamble they chose from that set would be resolved with a randomizing device and they would be paid accordingly.

A sample of the decision screen for gains is shown in Figure 5. As can be seen, the information on probabilities, gains, and the trade-offs was presented in three different ways on each decision screen. First, on the top of the screen, the budget is explained verbally. In this example, we state the maximum gain is $20 and each 1 percentage point increase in p will reduce gains by $0.25, that is $(r_1, r_2, m) = (0.25, 1, 20)$. Second, as the subject moved the slider to the right the green pie got larger, indicating a greater likelihood of winning, while the green bar got shorter, showing a smaller prize. Moving the slider left and right gives an interactive display of the rate of trade-offs between probability and reward. Probabilities were in presented integer units ("44 out of 100" in Figure 5, for instance) while marginal gains and losses were reported to pennies (for example, $16.67). Third, the bottom of the screen describes exactly what gamble is highlighted in the circle and bar above. Subjects are told to select the option they like *most.*

The decision screen for losses is similar, as seen in Figure 6, however colors switched from green to red. To further accentuate the difference between gains and losses, the red bar

[9] The experimental instructions stressed that the $20 endowment was already earned, and that choices could result in a "gain" or "loss" relative to this endowment. Thus, we perform analysis assuming that subjects care about the gain or loss from their choice rather than relative to some unobserved expectation. The analysis confirms this as a meaningful approach.

[10] In a pilot study we completely randomized presentation of the choices between the gain and loss sets. Feedback from the subjects was that flipping between losses and gains led to mistakes, such as reporting choices on losses as if they were for gains. We decided at that point to group all of the losses and all of the gains together. Checking for order effects, we found no significant differences due to order.

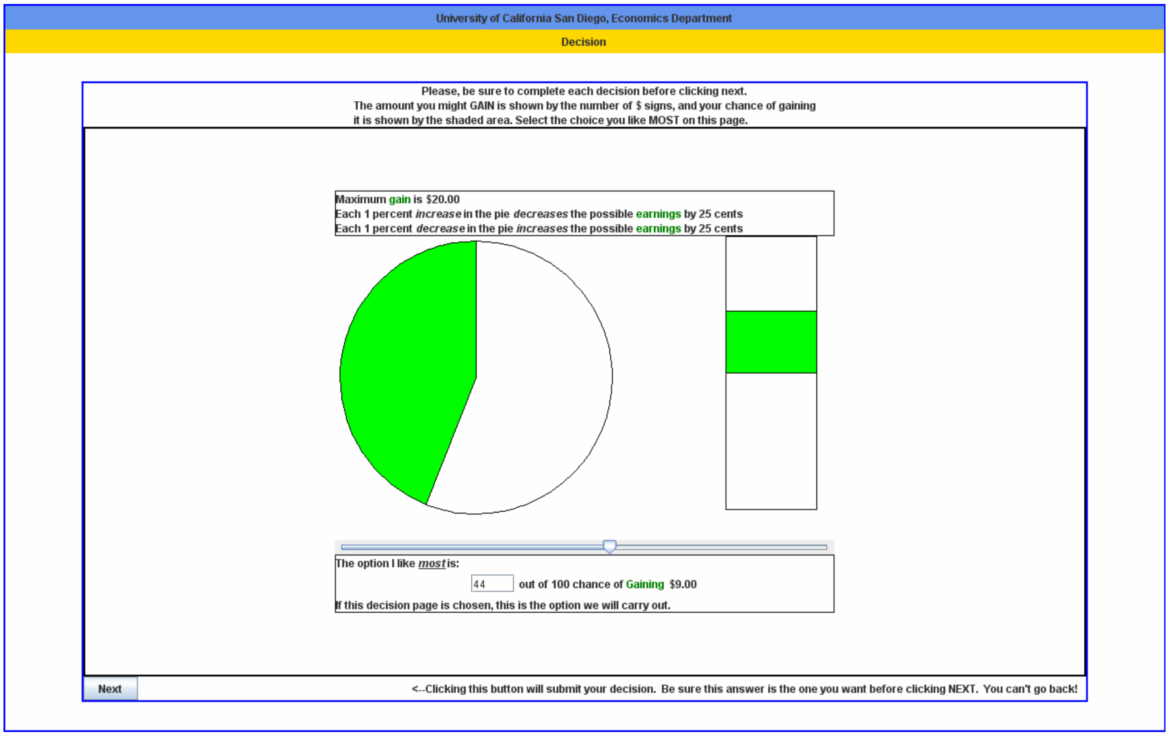


Figure 5: Decision Screen for Gains

(losses) grows down from zero while the green bar for gains (in Figure 5) grows up from zero.

For each budget set over losses, subjects were told to select the gamble they like *least* and that this gamble will definitely *not* be selected to be played – while any of the other possible gambles were equally likely to be selected, if this were the budget set chosen for payment. To make the minimization decision more meaningful, we also told subjects that we would eliminate the gamble they choose plus the two to the left and the two to the right. So in the example in Figure 6 the subject is choosing "11 out of 100 chance of losing $10.50" as the option preferred least, and is told that we will choose another gamble available on this page, but excluding those with chances 9, 10, 11, 12, and 13 out of 100.

Subjects were presented extensive instructions. These included examples of possible choices and outcomes, as well as 6 "quiz" questions that tested their understanding of the instructions. Subjects could not move forward until all participants answered all quiz questions correctly. A copy of the instructions can be found in the appendix to this paper, and a Java™ applet illustrating the dynamic choice screens can be found on the authors' websites.

We also built in checks to prevent subjects from seeing their choices as part of a portfolio in which they balanced risks of losses with risks of gains. First, all choice screens were presented one at a time, and in random order (except that gains and losses were all grouped together) and subjects did not see a "master list" of choices. Second, subjects were not able to go back and forth between budget sets and alter previous choices during the main part of the experiment. Third, we tested for the presence of portfolio-rebalancing by adding two more surprise stages. In particular, after completing all 28 decisions we publicly flipped

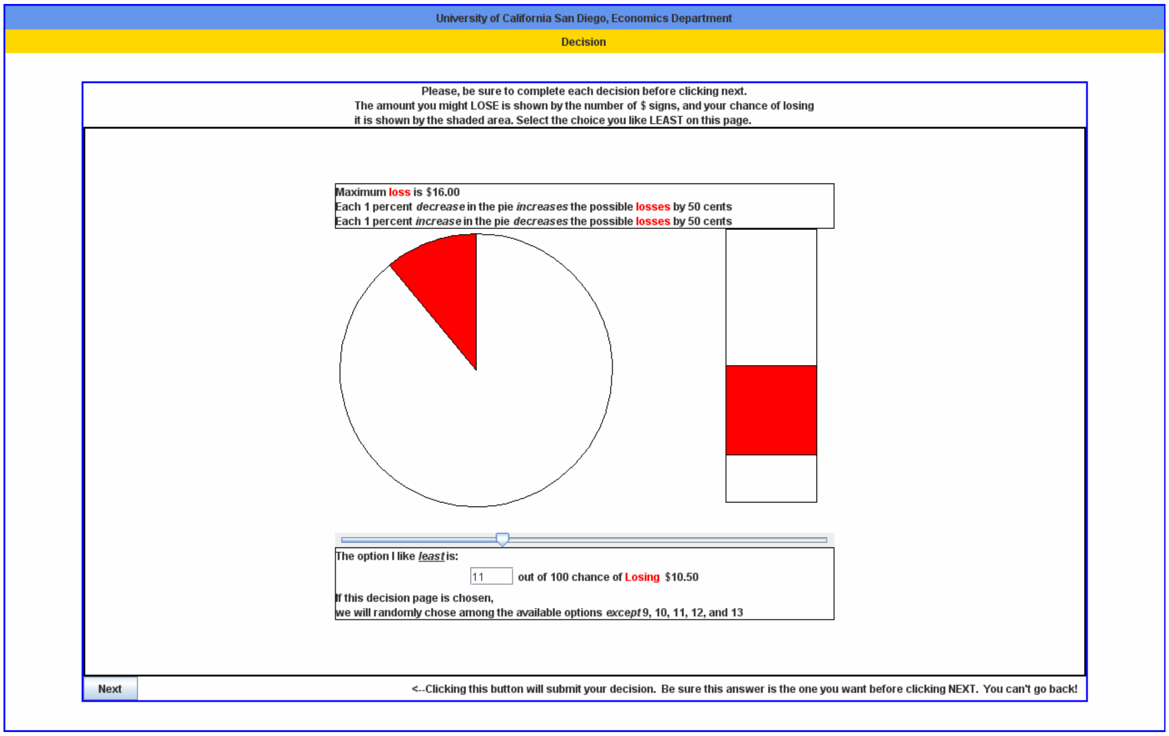


Figure 6: Decison Screen for Losses

a coin to determine whether to pay on gains or losses. We then entered "stage 2," where subjects could review all of their choices for the selected set (gains or losses), and make any revisions they pleased. When stage 2 was complete, we drew a number from 1 to 14 to determine which choice number we would use. This began stage 3 where we allowed subjects to alter their decision on the chosen budget set. Finally, we rolled two ten-sided die to determine payoffs. Since the subjects had no idea stages 2 and 3 were coming, if they were treating each decision as an independent choice rather than as a portfolio problem, then we should see no revisions in stages 2 and 3. In the gains domain 42 of the 88 subjects changed at least one choice between stages one and three. However, the average number of changes was only 3.43 of 14 possible and the average size of the absolute value of the changes was only 1.86 percentage points. In the loss domain 35 subjects made changes, with averages of 2.37 changes and 2.33 percentage points. Seeing only trivial changes, we discount any rebalancing effects and focus on the stage 1 results in the analysis below.

4 Aggregate Results

Figure 7 (left panel) shows the choice sets offered for gains and the average gamble selected on each. Selecting x as the numeraire, so $r_2 = 1$, then the set of relative prices, in cents, of probability is $r_1 \in \{6.25, 8.33, 16.67, 12.5, 25, 50, 100\}$. The set of normalized "incomes" is $m \in \{4, 8, 12, 16, 20, 24\}$. Figure 7 (right panel) shows the same information for losses.

The average choices across subjects is well organized. There are no violations of revealed

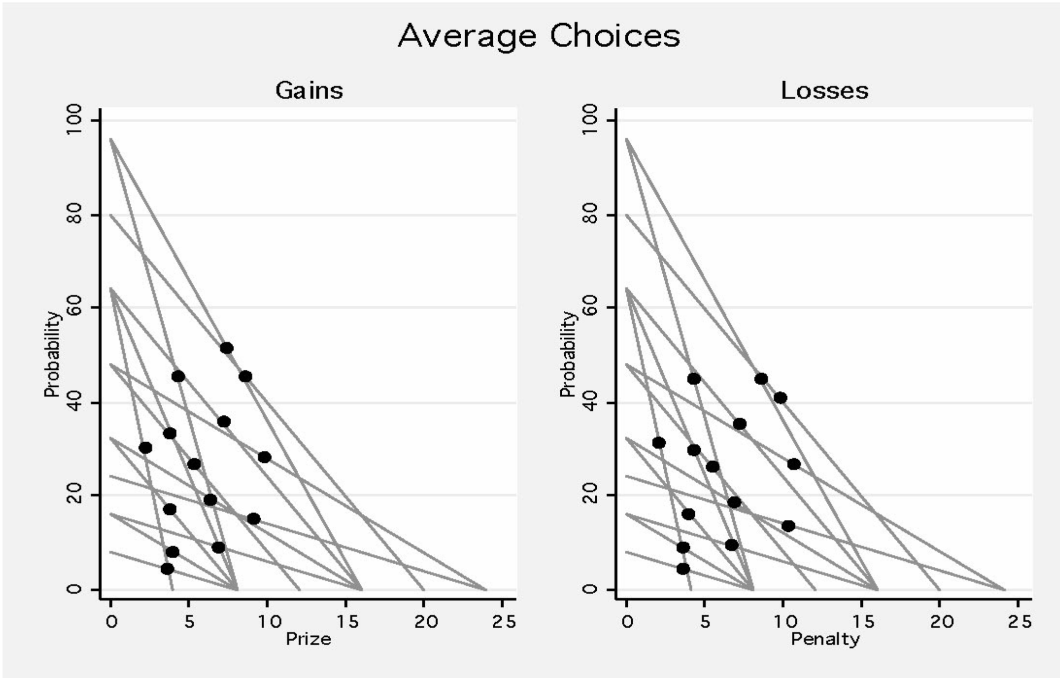


Figure 7: Average choices show no violatations of GARP or GARA

preference for either average gains or losses. Both graphs indicate choices consistent with downward sloping demands and with treating both p and x as normal goods. While this is encouraging, it is not particularly meaningful. The next section examines the data at an individual level and shows how it can be used to test specific hypotheses about risk behavior.

5 Testing Question 1: GARP and GARA

Here we ask whether choices are consistent with some well-behaved preference ordering by testing adherence to the axioms of revealed preference. We begin with a brief description of the tests.

5.1 Defining GARP and GARA

A bundle A is *directly revealed preferred* to a bundle B if B was available when A was chosen, written $A\mathbf{P}^dB$. If B is strictly within the budget set, then we say A is *strictly directly revealed preferred* to B. Finally, we say a bundle A is *revealed preferred* to a bundle Z if there exists a chain of directly revealed preferred comparisons, $A\mathbf{P}^dB, B\mathbf{P}^dC, ..., Y\mathbf{P}^dZ$, connecting A to Z. The axiom to test is GARP:

Definition 1 Generalized Axiom of Revealed Preference (GARP): *If A is revealed preferred to B, then B is never strictly directly revealed preferred to A.*

When we ask our subjects to minimize utility over budgets of bads, as we did with losses, the revealed aversiveness axioms mirror those of revealed preference.[11] Suppose $A', B', ...Z'$ are all bads. A bundle A' is *directly revealed more aversive* than a bundle B' if B' was available when A' was chosen as most aversive, written $A'\mathbf{A}^dB'$, and A' *strictly directly revealed more aversive* than B' if B' was strictly within the budget set when A' was chosen most aversive. Finally A' is *revealed more aversive than* B' if there is a chain of direct comparison such that $A'\mathbf{A}^dB', B'\mathbf{A}^dC', ..., Y'\mathbf{A}^dZ'$.We then apply this test:

Definition 2 Generalized Axiom of Revealed Aversiveness (GARA): *If A' is revealed more aversive than B', then B' is never strictly directly revealed more aversive than A'.*

A method devised by Varian (1982) for adding up the violations of revealed preference is to count the number of budgets involved in one or more violations. Thus, if $A\mathbf{P}^dB$ and $B\mathbf{P}^dA$, then we have a single violation of revealed preference, but two budgets are involved. Similarly with aversiveness.

If there is a violation of revealed preference or aversiveness, it may be because of a failure of rationality, or because of "errors" in choice. Nonparametric analysis therefore often allows some tolerance for "small" violations that would still be acceptable under the null hypothesis of rational choice, just as one might add a random error term to a parametric estimation. The common tool to measure this is the Afriat Critical Cost Efficiency Index (Afriat, 1972), or CCEI for short. Intuitively, the CCEI measures the fraction of a budget "wasted" by not satisfying GARP or GARA. For instance, if CCEI equals 0.80 it means an individual could have purchased a revealed preferred bundle at 80% of what she actually spent.[12] CCEI equal to 1 means there are no violations of GARP. It is generally accepted that values of the CCEI above 0.95 should be seen as "small"errors.

5.2 Power of GARP and GARA tests

If subjects fail to violate GARP or GARA, it is natural to ask how powerful the test was at uncovering irrationality if it were there. First we calculate Bronars' (1987) index with a simulation of 10,000 synthetic subjects, each making random choices from our budgets with uniform probabilities of each possible choice. These synthetic subjects average 6.58 violations and an CCEI of 0.803. We also use a bootstrapping test (Andreoni and Harbaugh, 2009) which samples from the distribution of actual choices over each budget set. This simulation is designed to incorporate the available information on average preferences into the power calculation and is therefore different between losses and gains. The simulation with our

[11] To our knowledge, this is the first expression of the notion of Revealed Aversion and the statement of any axioms in the domain of preferences. Nonetheless, the application to revealed cost minimization is nearly identical to these ideas and has been extensively developed by Varian (1984).

[12] The CCEI is 1 minus the proportion by which the budget sets need to be moved towards the origin until all intransitivities disappear. Specifically, for $0 < e \leq 1$ we say $A\mathbf{R}^d(e)B$ if $A\mathbf{R}^d(eB)$, where $\mathbf{R}^d$ refers to either $\mathbf{P}^d$ or $\mathbf{A}^d$. Then find maximum value of e such that, using $\mathbf{R}^d(e)$ rather than $\mathbf{R}^d$, there are no violations of revealed preference in the data. This e is the CCEI. For more detail on this, see Varian (1982, 1991), and for applications to experimental data see Andreoni and Miller (2002), Harbaugh, Krause, and Berry (2001), Choi, et al. (2007), or Andreoni (2008).

budget sets and the distributions from our sample produce 2.95 GARP violations and an CCEI of 0.933 over gains, and 6.58 GARA violations and an CCEI of 0.751 over losses. In short, our subjects had ample opportunity to violate GARP and GARA in this protocol.

5.3 Results

Table 1 lists the actual number of violations of GARP and GARA for our sample for both gains and losses. Look first at gains. This data is supportive of a general coherence with GARP: 57 subjects (65%) have no violations, and 85% of subjects have either 0 or 2 budgets involved in violations of GARP. The data on losses is less encouraging, although still supportive of rationality for most subjects. Thirty-eight subjects (43%) have no violations, and 59% have either 0 or 2 budgets in violation of GARA.

Table 1
Revealed Preference Violations for Gains, and Revealed Aversive Violations for Losses

Number of Budgets Involved in Revealed Preference Violations	Gains: GARP Violations	Losses: GARA Violations
0	57	32
2	18	14
3	1	9
4	3	7
5	4	3
6	2	6
7		4
8	1	2
9		1
10	1	4
11		3
12	1	1
13		1
14		1
Total	88	88

Table 2 shows a similar pattern on average. For gains, there are 1.284 GARP violations on average. The average CCEI is 0.976 (s.d. 0.058), which is quite close to 1, with a small standard deviation, and highly supportive of a model of rational choice. For losses the average number of budgets in violation of GARA is 3.398, which is more than twice that of gains. The CCEI for losses averages 0.878 which is below the critical value of 0.95, but the standard deviation is 0.158, indicating a great deal of variation across individuals. This implies that the data on gains is far more consistent within subjects than that on losses.

It is impossible to know whether this is because preferences over losses are less rational, or whether our protocol on losses—asking people to choose their least preferred outcome—was a more difficult or confusing task. Informally, subjects said after the experiment that the minimization was far more difficult, supporting the latter hypothesis, with some even stating that, despite the many cues in the protocol, when they switched from gains to losses they absentmindedly started by maximizing rather than minimizing. Formally, the correlation of the CCEI on gains with that on losses is $\rho = 0.019$, which means there is no tendency for those who have large violations in gains to also have large violations in losses. This supports the view that the task on losses was simply more difficult and, as a result, the data is more noisy on losses.

Table 2
Revealed Preference Violations
and Afriat Critical Cost Efficiency Index
by Gain or Loss*

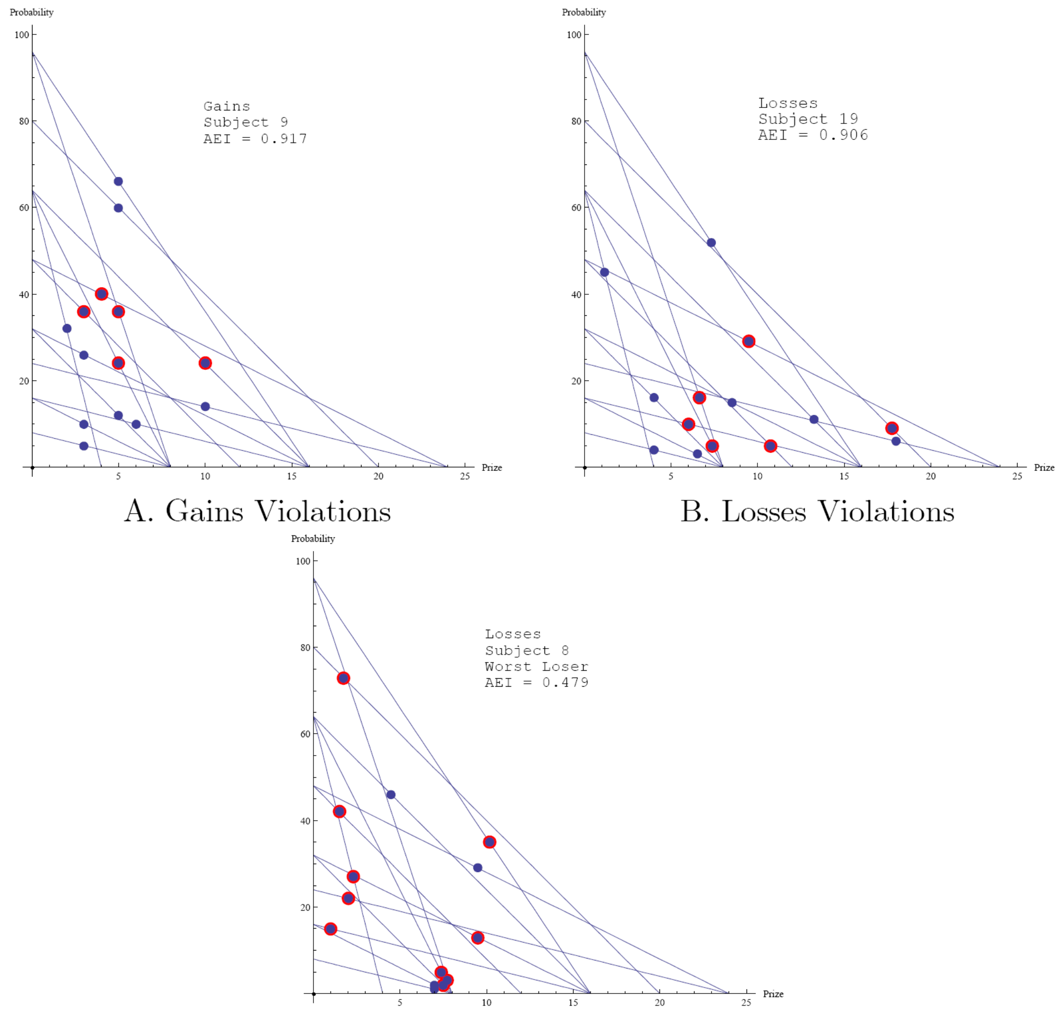
	Stage 1	
	Mean	St. Dev.
Gains:		
Ave GARP Violations:	1.284	2.309
Ave Gain CCEI:	0.976	0.058
Losses:		
Ave. GARA Violations:	3.398	3.716
Ave Loss CCEI:	0.878	0.158

*88 observations reported in each row.

Figure 8 illustrates a typical violation in gains (panel A) and losses (panel B), with the offending choices highlighted in red. Panel C shows the individual with the lowest CCEI in the sample. This subject appears to be one of those who was mistakenly maximizing for a few choices rather than minimizing—she selected corner solutions on a number of budgets, and hence generated large violations of GARA.

The problem of individuals maximizing rather than minimizing on losses, as in panel C above, seems to explain a large fraction of the difference between gains and losses. In gains there were only three occasions in which a subject chose an allocation with an expected value of zero, while in losses that number was 80. Such errors also need not violate GARA, as shown in Figure 9 which plots the choices of subject 57, who does not violate revealed preferences if we believe the person maximizes utility on gains but minimizes it on losses. Such an implausible conclusion points to confusion on the task rather than true preferences.

We conclude that choices on gains are consistent with a model of rational choice. The data on losses is also suggestive of general coherence to rational choice, although there is a great deal more heterogeneity in violations and a significant number of subjects who violate rationality. Our best guess is that this asymmetry is due to the unusual protocol in losses, that is, asking people to report their least rather than their most preferred choice.



A. Gains Violations

B. Losses Violations

C. Worst Violations

Figure 8: Examples of Violations of GARP and GARA

6 Testing Question 2: Prospect Theory Asymmetry

Our framework allows three perspectives for examining prospect theory asymmetry. First is a nonparametric look at risk aversion and risk loving over gains and losses. Second is a parametric analysis of individual data. Third is parametric analysis on aggregate data. Our analysis is supportive of prospect theory loss aversion for about one third of individuals. These individuals have a strong enough influence on the aggregate data to make risk aversion on gains and risk loving on losses the best fitting model overall. However, when we trim the sample to exclude choices that are very likely the result of confusion, we find subjects on average are only risk neutral on losses.

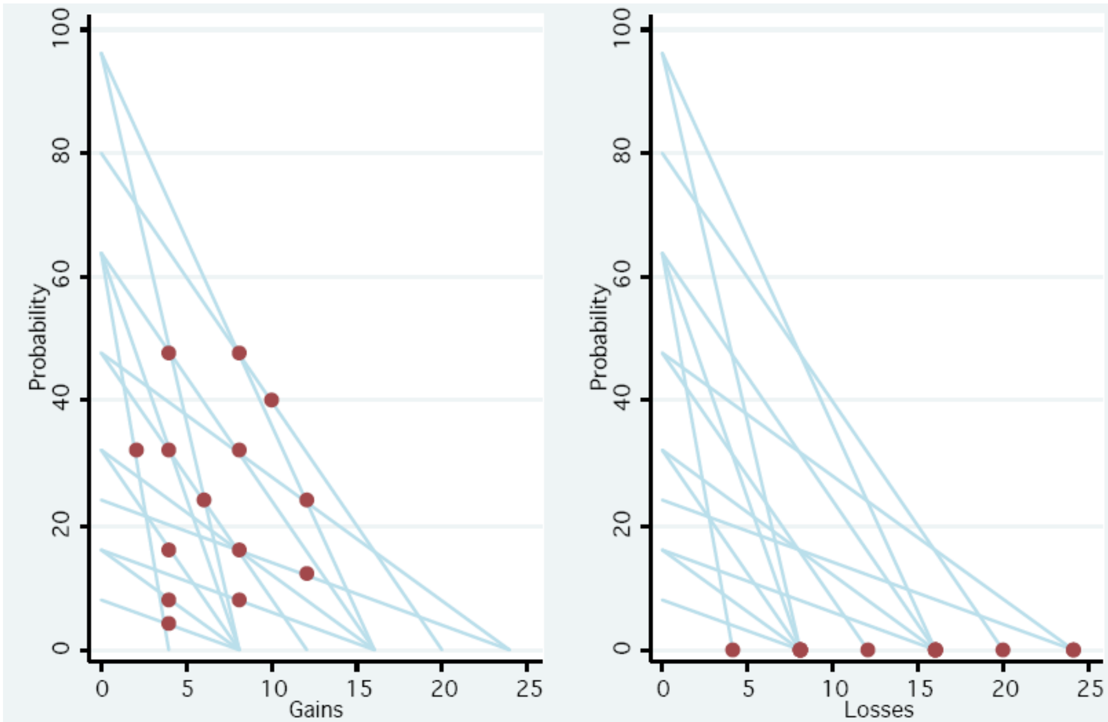


Figure 9: Subject 57 who behaves as if maximizing utility on gains while minimizing utility on losses.

6.1 Individual level: Nonparametric Tests

As shown in Figure 2, a choice is risk averse if x is less than the expected value maximizing x for gains, and $|x|$ is higher than the expected value minimizing $|x|$ for losses. Here we calculate the fraction of choices in gains and losses that are risk averse or risk loving. We will say a person exhibits risk aversion if 50% or more choices are risk averse, and risk loving if more than 50% are risk loving. Figure 10 plots these proportions for every subject, divided into the four possible combinations.

Figure 10 reveals that 60.2% of subjects are risk averse on gains, 56.8% are risk loving in losses, but only 34.1% are both risk averse on gains and risk loving on losses. 26.1% are risk averse everywhere and 22.7% are risk loving everywhere, while 17% are inexplicable risk loving on gains and risk averse on losses.

This test shows evidence of the risk preferences proposed by Kahneman and Tversky (1979), with about one third of subjects fitting their model of both concave preferences for gains and convex preferences for losses.[13]

[13]Stating an alternative hypothesis in which subjects choose randomly, one would expect about 25% to fit in each of the four quadrants showen in Figure 10 A test of the random assignment to the four "types"cannot reject random assignment to the four cells ($\chi^2[3] = 5.36$), hence, although the propect theory type is the most frequent, the difference from the others is not significant.

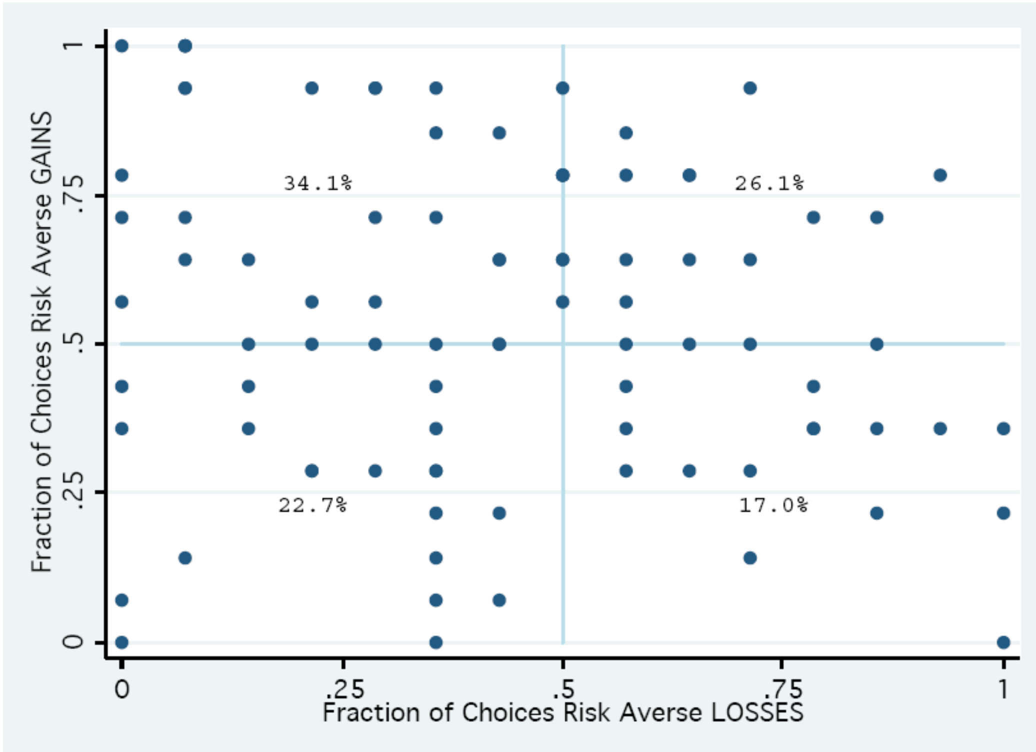


Figure 10: Revealed Risk Aversion and Risk Loving by Subject

6.2 Individual Level Test: Parametric Analysis

Since there are 14 choices for gains and 14 choices for losses, there are enough observations to estimate a simple utility function for each subject over each domain. Following the literature, we estimate the following:[14]

$$u(x) = \begin{cases} x^{\alpha} & \text{if} \quad 1 < x \\ x & \text{if} \quad -1 \leq x \leq 1 \\ -(-x)^{\beta} & \text{if} \quad x < -1 \end{cases} . \tag{8}$$

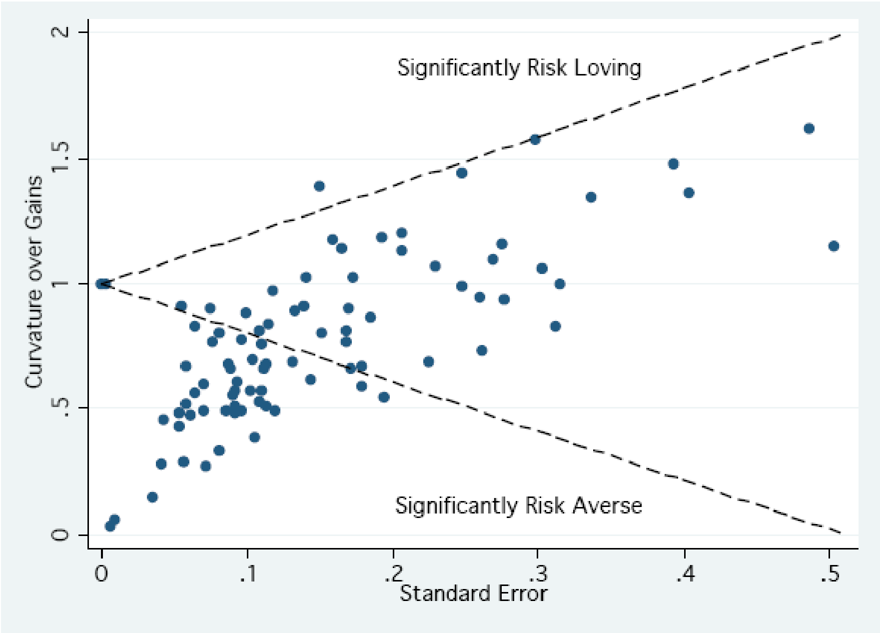
The restrictions for risk aversion are that $0 < \alpha < 1$, and $\beta > 1$. Risk loving follows from $\alpha > 1$ and $0 \leq \beta < 1$.

We estimate parameters by first solving for the demand for x. For gains $x(m, r_2) = [\alpha/(1+\alpha)]m/r_2 = am/r_2$, where $a = \alpha/(1+\alpha)$. We then estimate a by ordinary least squares (OLS), say $\hat{a}$, solve $\hat{\alpha} = \hat{a}/(1-\hat{a})$ and find the standard error using the delta method. We proceeded similarly for estimating $\hat{\beta}$. Results are shown in Figure 11.

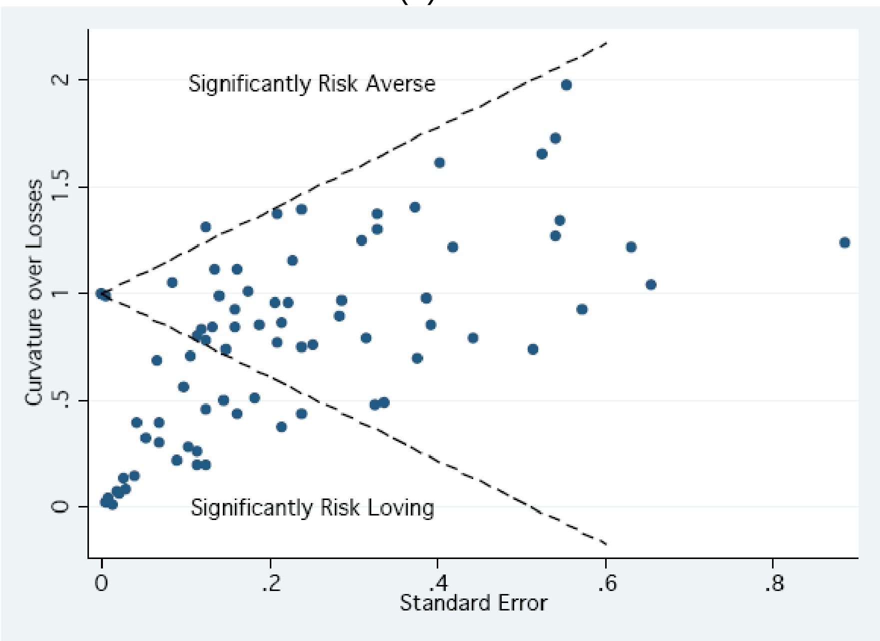
Looking first to gains, we find that 48.9% of the subjects have values of $\hat{\alpha}$ significantly below 1, indicating risk aversion on gains, and only 1 shows significant risk loving. 47.7% have α insignificantly different from 1.

Over losses, we find 29.5% of subjects with β significantly below 1 for risk loving, and again only 1 subject with β significantly greater than 1 for risk aversion. 67.0% of subjects

[14] Note the function is linearized for $-1 \leq x \leq 1$ since the slope of x^{α} nears infinity as x nears 0.



(a) Gains



(b) Losses

Figure 11: Estimates and standard errors of (a) $\hat{\alpha}$ and (b) $\hat{\beta}$ by subject. Dashed lines indicate ± 1.96 standard deviations from risk neutrality.

have β insignificantly different from 1.

Considering gains and losses together, 17% of subjects have prospect-theory preferences that are significantly risk averse on gains and risk loving on losses. Only one person is significantly risk averse everywhere, and none show risk loving everywhere.

Relative to the nonparametric tests, this parametric analysis decreases the fraction of subjects with prospect theory loss aversion by half. Nonetheless, it remains the category that is most prominently supported by significant coefficients. Because we cannot precisely estimate preference parameters for such a large fraction of our subjects, however, we are unable to reject standard assumptions of expected utility for 82% of them.

Table 3
Estimates of Aggregate Utility Function

	Full Sample	Limited Sample
Parameter Estimates:		
$\hat{\alpha}$	0.738	0.739
	(0.040)	(0.040)
$\hat{\beta}$	0.840	1.014
	(0.075)	(0.074)
F-stats for test of:		
$\hat{\alpha} = \hat{\beta}$	1.8	12.23
	($p = 0.18$)	($p < 0.01$)
$\hat{\alpha} = 1$	43.96	43.4
	($p < 0.01$)	($p < 0.01$)
$\hat{\beta} = 1$	4.54	0.03
	($p = 0.04$)	($p = 0.85$)
$\hat{\alpha} = \hat{\beta} = 1$	22.15	22.53
	($p < 0.01$)	($p < 0.01$)
Observations	2464	2175
Clusters	88	88

Note: Standard errors clustered at individual in parentheses. Restricted sample excludes observations with expected value of zero, which eliminated 3 gains data points and 80 losses.

6.3 Aggregate Data: Parametric Analysis

In this analysis we pool the data and estimate (8) with errors clustered at the individual level. We conduct our estimates on two different samples. First, we consider all 88 subjects in the analysis. Next we try to account for our own failings as experimenters and exclude the individual choices in which the expected value is zero. The premise for this is that selecting an expected value of zero would never be the result of optimizing behavior and thus is likely to be due to confusion within the subject. Of the 2464 observations, this drops 3 observations from gains and 80 from losses.

The results of these regressions are reported in Table 3. Using the full sample for gains we find $\alpha = 0.738$ and significant evidence of risk aversion. On losses, we find $\beta = 0.840$, and significant evidence of risk loving. Although it lacks statistical significance (we cannot reject that $\hat{\alpha} = \hat{\beta}$), the point estimates indicate there is a kink at 0 gains and losses, with slightly greater concavity over gains than convexity over losses, as posited by prospect theory.

The limited sample gives a somewhat different picture. Gains stay largely the same, while for losses the estimate is $\beta = 1.014$ (s.e. 0.074), indicating risk neutrality. This suggests that the subjects who may have been confused were driving the aggregate estimates of risk loving in losses. Nonetheless, the analysis resorts the first order effect of loss aversion, that is, utility is significantly steeper in losses.[15]

7 Testing Questions 3 and 4: The Independence Axiom vs. Probability Weighting

Figure 12 shows the offer curves for the average choices for both gains and losses. To the naked eye, these curves seem either vertical, which would be consistent with the independence axiom, or perhaps slightly upward sloping, which would contradict both independence and probability weighting.[16] Next we explore the statistical and economic significance of possible deviations from the independence axiom.

7.1 Statistical and Economic Significance

To test whether the slopes of the offer curves deviate significantly from vertical (infinity), imagine inverting the axes in Figure 12 and asking whether the slope of the offer curves are significantly different from zero, with a positive slope being consistent with probability weighting. To do this, first normalize the budget constraint so that the x is the numeraire, that is $r'p + x = m'$, where $r' = r_1/r_2$ and $m' = m/r_2$. Then regressing x on the price r' with fixed effects for m' can tell us whether the coefficient on r' is significantly different from zero, and interactions with r' and the fixed effects can tell us whether there is a difference across

[15] One will note that the estimate of α in this analysis is similar to that found by Andreoni and Sprenger (2009a,b), who use different measures but restrict estimation to situations involving only risk, and that the value of 0.738 is much less extreme risk aversion than found by many other researchers. For instance, in the auction literature, mention is made of 'square root utility' where $\alpha \approx 0.5$. Holt and Laury (2002) discuss several relevant willingness to pay results from the auction literature in line with this value. Nonetheless, this less extreme risk aversion is still concave enough to suffer from the criticisms of Rabin (2000a). As seen in the answer to question 2, however, support for reference dependence, one of the solutions envisioned to Rabin's critique (Rabin 2000b), is also suggested in our data.

[16] With only 11 points across three budgets, it is artful at best to conjecture about individual offer curves. Nonetheless, one sesible criteria on individuals is the following: Look at differences between choices on adjacent budgets with the same m. Count whether these changes are positive, negative, or zero. Categorize a change as zero if the two values of x are within 0.5. This accounts for the granularity of choice data. Then say an offer curve is upward (downward) sloping if there are two more upward (downward) changes than downward (upward). Then for gains we would say there are 21% upward sloping, 1% downward sloping, and 78% vertical offer curves. On losses, the respective percentages would be 26%, 5% and 69%.

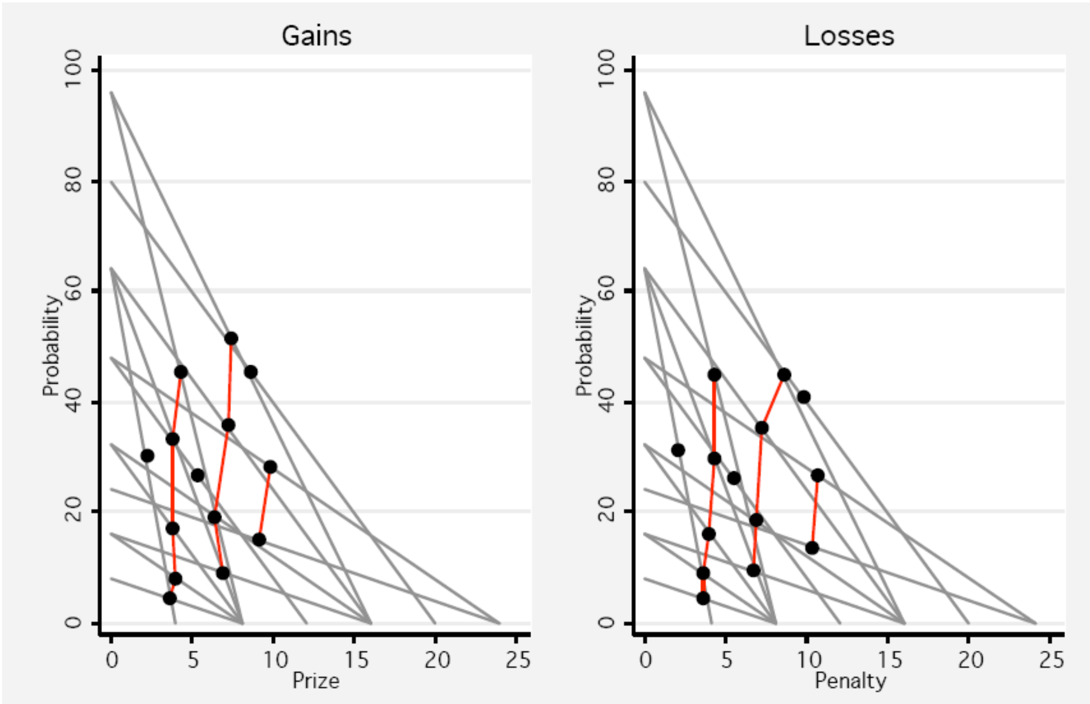


Figure 12: The independence axiom implies offer curves are vertical, while probability weighting predicts they slope back.

m's. These regressions are presented in Table 4 for the 11 budgets that allow this test.[17]

Table 4 shows a clear rejection of probability weighting and mixed results on the independence axiom. Looking just at the effect of r' combined across all three offer curves, the coefficient is negative and significant, which contradicts both assumptions. Looking across offer curves for different incomes, however, only that for $m' = 8$ finds a coefficient significantly different from zero on gains, although $m' = 8$ and 16 find significance for losses.

[17] We also ran identical regressions for the restricted sample as defined in Table 3 above, with very similar results.

Table 4
Test of vertical offer curves: Fixed effects regression of prize x on relative price of p, $r' = r_1/r_2$, holding constant real income $m' = m/r_2$. A significant positive slope on price of p is consistent with probability weighting, while a zero coefficient is consistent with the independence axiom. ‡

	Gains	Gains	Losses	Losses
r' = price of p	**−0.632**		**−1.124**	
	(0.238)		(0.371)	
$(m' = 8) \times r'$		**−0.493**		**−0.87**
		(0.183)		(0.309)
$(m' = 16) \times r'$		−0.620		*−1.597*
		(0.366)		(0.695)
$(m' = 24) \times r'$		−1.307		−0.659
		(1.084)		(1.705)
$m' = 8$	**4.111**	**4.057**	**4.343**	**4.246**
	(0.148)	(0.134)	(0.200)	(0.197)
$m' = 16$	**7.232**	**7.226**	**7.845**	**8.072**
	(0.158)	(0.208)	(0.273)	(0.393)
$m' = 24$	**9.949**	**10.455**	**11.314**	**10.966**
	(0.301)	(0.829)	(0.562)	(1.401)
Observations	968	968	968	968
Clusters	88	88	88	0.88

‡Clustered errors in parentheses, **bold** is $p < 0.01$, and *italic* is $p < 0.05$.

While these coefficients are statistically significant, perhaps the larger question is are they economically significant? What, for instance, does the coefficient of −0.632 mean in practical terms? Table 5 gives some insight into this question. Here we look at the minimum and maximum price on each offer curve shown in Figure 12, and the associated quantity. On the $m' = 8$ offer curve, for instance, price of p decreases by 1100%, and the demand for x goes up by 18%. While the change in x is statistically significant, the "gross elasticity" is only $\epsilon = -0.015$. Similar results can be seen for m' of 16 and 24. Thus, while our measurements are precise enough to find statistical significance, it is more challenging to argue that the deviation from the independence axiom is meaningful in an economic sense.[18]

[18] We understand that one could also use this same argument to claim that the economic significance of the difference with probability weighting is also not meaningful. We appeal here to Occam's razor that the simpler model should take precidence as the null hypothesis in this test, thus take this data to be more supportive of the independence axiom than probability weighting.

Table 5.
Testing the Economic Significance of imposing the Independence Axiom on Gains: Absolute change is significant for $m' = 8, 16$, but Gross Elasticity is economically small.

Real Income, $m' = m/r_2$:	$m' = 8$	$m' = 16$	$m' = 24$
Real price of p, r_1/r_2 :			
Maximum	1	1	1
Minimum	1/12	1/8	1/2
Percent change	1100%	700%	100%
A. Gains			
Mean choice of x :			
At maximum price	\$3.53	\$6.83	\$9.15
At minimum price	\$4.22	\$7.42	\$9.80
Percent change	−16%	−8%	−7%
Comparisons:			
Diff. of mean x, t-test	3.07	1.42	0.97
Gross Elasticity*	−0.015	−0.016	−0.067
B. Losses			
Mean choice of x :			
At maximum price	3.52	6.73	10.31
At minimum price	4.27	8.51	10.64
Percent change	-17%	-21%	-3%
Comparisons:			
Diff. of mean x, t-test	2.10	2.47	0.31
Gross Elasticity*	-0.016	-0.042	-0.031

*Gross Elasticity is defined as $(\Delta x/x)/(\Delta p/p)$.

This test allows one clear conclusion: there is no evidence to support probability weighting where low probabilities are overweighted and high probabilities underweighted as an alternative to the independence axiom.[19] There is some evidence to suggest that offer curves in Figure 12, while very near vertical, have a significant positive slopes, in contradiction to the independence axiom. However we would argue that the economic significance of this deviation from vertical leaves little room for a theorist's creativity to flourish. Stated differently, the data does not provide a compelling case for rejecting the independence axiom.

8 Testing Assumption 5: CRRA Utility

Since CRRA preferences are typically constructed in the domain of gains, we will restrict our analysis to the data on gains. Recall, CRRA utility is identical to Cobb-Douglas preferences. There are two main restrictions that capture Cobb-Douglas preferences, and both are testable

[19] We also reproduced Table 5 using the restricted sample discussed above. Results for gains are virtually unchanged, but those for losses become more supportive of the independence axiom. The t-stats all become insignificant $(1.39, 1.60, 0.16$, respectively) and the gross elasticities are closer to zero by half or more $(-0.008, -0.022, -0.013$, respectively)

with the data. The first restriction is that demands should have constant budget shares. In particular $r_1p/m = 1/(1+\alpha)$ and $r_2x/m = \alpha/(1+\alpha)$. Let $s_{it} = r_{2t}x_t/m_t$ be the budget share for x by person i on budget t. As above, let $m'_t = m_t/r_2$ and $r'_t = r_{1t}/r_{2t}$. Then the regression equation $s_{it} = \theta_o + \theta_1 m'_t + \theta_2 r'_t + \epsilon_{it}$ should yield values of $\theta_1 = \theta_2 = 0$ under the null hypothesis of CRRA preferences. Moreover, the value of θ_0 should be consistent with the estimate of $\hat{\alpha}$ from the aggregate analysis presented in Table 3, in particular $\theta_0 = \hat{\alpha}/(1+\hat{\alpha}) = 0.42$.

Results of this regression are as predicted. First, $\hat{\theta}_0 = 0.422$, ($s.e. = 0.009$). Second, the coefficients $\hat{\theta}_1$ and $\hat{\theta}_2$ on both are very small and extremely precisely estimated: $\hat{\theta}_1 = -0.003$ ($s.e. = .0002$), and $\hat{\theta}_2 = 0.094$ ($s.e. = 0.0013$). Although the estimates of both parameters are significantly different from zero statistically, the economic significance of each point estimate is small. Doubling the price of p from 1 to 2, for instance, increases the budget share of x by less than 1 percent.[20]

A second restriction of Cobb-Douglas is that demands reflect a constant elasticity of substitution, σ, and that $\sigma = -1$. We test this by regressing $\ln(p/x)$ on $\ln(r_1/r_2)$. The coefficient on $\ln(r_1/r_2)$ is our estimate of σ. Doing so yields an coefficient $\hat{\sigma} = -1.063$ ($s.e. = 0.019$). Again, the coefficient is negligibly different from 1 but is precisely estimated, leading to a rejection of the hypothesis that $\sigma = 1$. Still, the estimate of σ is likely close enough to 1 to make most economists comfortable with an assumption of CRRA preferences.[21]

9 Summary and Conclusion

The experimental literature on expected utility and its alternatives has used a variety of methods and produced a mix of results. Direct measures of preferences or tests of assumptions are often based on discrete choices among a limited set of dissimilar alternatives. Moreover, one of the alternatives is often a sure thing, which confounds choices over risk with the Allais certainty effect. Despite hundreds of papers, the literature has failed to cohere around a single model or measurement tool. We use a new method to measure preferences that is simple, continuous, and involves only risky choices. Our method allows clear and clean tests of central questions about the standard model of expected utility and alternatives to it. The resulting data as a whole is supportive of the model of expected utility, including the independence axiom, and it rejects some prominent alternatives, such as probability weighting. Aggregate analysis also supports the prospect theory notion of loss aversion around a reference point of zero gains and losses, but not the assumption of risk loving over losses.

Our method allows subjects to trade-off higher rewards for lower probabilities, by choosing their optimal (p, x) along a downward sloping budget $r_1p + r_2x = m$. For losses, they choose their least favorite options of $(p, |x|)$ along the budget $r_1p + r_2|x| = m$. We elicit choices for each subject from budgets with differing incomes and slopes.

Our analysis starts with the simple assumption that people maximize some general utility function over the probability and payoff of a gamble $U = U(p, x)$. We use this method to test five key questions on preferences for risk. First, we show choices over gains largely obey the

[20] This regression included all 88 subjects, however regressions with the limited sample were nearly identical. To check the robustness of this regression we also added $(r')^2$ to the regression equation, but the results do not change the inferences we present here.

[21] This regression is also for the full sample. For the restricted sample the results are again nearly identical.

revealed preference axiom of the rational choice model. Losses, by contrast, generate much noisier behavior, a large part of which is likely attributable to the novelty and complexity of the choice task on losses as we presented it. Second, the individual data show 17–34% of subjects meet the conditions of prospect theory, by being risk averse on gains and risk loving on losses. Aggregate analysis also supports the prospect theory notion of loss aversion around a reference point of zero gains and losses, but not the assumption of risk loving over losses. Our third and fourth results follow from the observation that demand for x is largely independent of the price of p. Deviations from this are economically negligible. This supports the independence axiom and rejects probability weighting, and is perhaps most striking and important findings of the paper. Fifth, we show that CRRA utility puts constraints on choices in our protocol that budget shares should be constant and the elasticity of substitution could be negative one. Average choices in our data deviate only slightly from these restrictions.

Overall our results support the standard assumptions of the expected utility when considering gains, including rationality and risk aversion. When losses are involved, choices are more noisy and prospect theory loss-aversion is necessary to explain the choices of a significant minority of subjects. Probability weighting, by contrast, is rejected for both gains and losses; instead the data favors the independence axiom. Finally, if a researcher would like to impose the simplification of CRRA utility, this likely comes at a small cost on average.

References

Afriat, S. "Efficiency Estimates of Productions Functions." *International Economic Review*, 1972, 13, pp. 568-598.

Allais, Maurice. "Fondements d'une théorie positive des choix comportant un risque et critique des postulats et axiomes de l'Ecole Américaine." *Colloques Internationaux du Centre National de la Recherche Scientifique Econométrie*, 1953, 40, pp.257-332.

Allais, Maurice. "Allais Paradox," in Steven N. Durlauf and Lawrence E. Blume, eds., The New Palgrave Dictionary of Economics, 2nd ed., Palgrave Macmillan, 2008.

Andreoni, James, "Giving Gifts to Groups: How Altruism Depends on the Number of Recipients." *Journal of Public Economics*, v. 91, September 2007, 1731-1749.

Andreoni, James and Harbaugh, William T. "Power Indices for Revealed Preference Tests." March 2006, manuscript.

Andreoni, James and John H. Miller, "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism." *Econometrica*, v. 70, no.2, March 2002, 737-753.

Andreoni, James, and Charles Sprenger, "Estimating Time Preferences from Convex Budgets." December 2009a. manuscript.

Andreoni, James, and Charles Sprenger, "Risk Preferences Are Not Time Preferences." December 2009b. manuscript.

Andreoni, James, and Charles Sprenger, "Certain and Uncertain Utility: The Allais Paradox and Five Decision Theory Phenomena." December 2009c. manuscript.

Bronars, Stephen G. "The power of Nonparametric Tests of Preference Maximization." *Econometrica*, 1987, 55, pp.693-698.

Choi, Syngjoo; Fisman, Raymond; Gale, Douglas and Kariv, Shachar. "Consistency and Heterogeneity of Individual Behavior under Uncertainty." *American Economic Review*, 2007, 97, pp.1921-1938.

Harbaugh, William T., Kate Krause, and Tim Berry. "GARP for Kids: On the Development of Rational Choice Behavior." *American Economics Review*, 2001, 91, pp. 1539-1545.

Harbaugh, William T., Kate Krause, and Lise Vesterlund. "The Fourfold Pattern of Risk Attitudes in Choice and Pricing Tasks." *Economic Journal.*, Forthcoming.

Harrison, Glenn W., Morten I. Lau , and E. Elisabet Rutström. "Estimating Risk Attitudes in Denmark: A Field Experiment." *Scandinavian Journal of Economics*, 2007, pp. 341-368.

Harrison, Glenn W. and E. Elisabet Rutström, "Risk Aversion in the Laboratory," *Research in Experimental Economics*, 2008, 12, pp. 41–196.

Holt, Charles A. and Laury, Susan K. "Risk aversion and incentive effects." *American Economic Review*, 2002, 92, pp.1644–1655.

Kahneman, Daniel and Amos Tversky. "Prospect theory: an analysis of decision under risk." *Econometrica*, 1979, 47, pp. 263–291.

Leland, Jonathan W. "Similarity Judgments in Choice under Uncertainty: A Reinterpretation of the Prediction of Regret Theory." *Management Science*,1998, XLIV, pp. 659–672.

Leland, Jonathan W. "Equilibrium Selection, Similarity Judgments and the 'Nothing to Gain/Nothing to Lose' Effect." manuscript 2008.

Neumann, John von and Morgenstern, Oskar. *Theory of Games and Economic Behavior*. Princeton, NJ. Princeton University Press, 1944.

Prelec, Drazen. "The Probability Weighting Function," *Econometrica*, May, 1998, Vol. 66, No. 3, pp. 497-527.

Quiggin, John. "A theory of anticipated utility." *Journal of Economic Behaviour and Organization*, 1982, 3, pp. 323–343.

Rabin, Matthew. "Risk aversion and expected-utility theory: a calibration theorem." *Econometrica*, 2000, 68, pp.1281–1292.

Rabin, Matthew, "Diminishing Marginal Utility of Wealth Cannot Explain Risk Aversion." in Daniel Kahneman and Amos Tversky, eds., *Choices, Values, and Frames*, New York: Cambridge University Press, 2000b, 202-208.

Rubinstein, Ariel, "Similarity and decision-making under risk (Is there a utility theory resolution to the Allais paradox?)". *Journal of Economic Theory*, 46(1), 1988, 145-153.

Tversky, Amos and Daniel Kahneman. "Advances in prospect theory: Cumulative representation of uncertainty." *Journal of Risk and Uncertainty*, 1992, 5, pp.297–323.

Tversky, Amos and Craig R. Fox, "Weighing Risk and Uncertainty." *Psychological Review*, 1995, 102 (2), pp. 269-283.

Varian, Hal R. "The Nonparametric Approach to Demand Analysis." *Econometrica*, 1982, 50, pp. 945-972.

Varian, Hal R."The nonparametric approach to production analysis." Econometrica: *Journal of the Econometric Society*, 1984

Varian, Hal R. "Goodness of Fit for Revealed Preferences Tests." University of Michigan CREST Working Paper Number 13, 1991.