

NBER WORKING PAPER SERIES

UNCERTAINTY EQUIVALENTS:
TESTING THE LIMITS OF THE INDEPENDENCE AXIOM

James Andreoni
Charles Sprenger

Working Paper 17342
<http://www.nber.org/papers/w17342>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2011

We acknowledge the generous support of the National Science Foundation Grant SES1024683 to Andreoni and Sprenger, and SES0962484 to Andreoni. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2011 by James Andreoni and Charles Sprenger. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Uncertainty Equivalents: Testing the Limits of the Independence Axiom
James Andreoni and Charles Sprenger
NBER Working Paper No. 17342
August 2011
JEL No. D81

ABSTRACT

There is convincing experimental evidence that Expected Utility fails, but when does it fail, how severely, and for what fraction of subjects? We explore these questions using a novel measure we call the uncertainty equivalent. We find Expected Utility performs well away from certainty, but fails near certainty for about 40% of subjects. Comparing non-Expected Utility theories, we strongly reject Prospect Theory probability weighting, we support disappointment aversion if amended to allow violations of stochastic dominance, but find the u-v model of a direct preference for certainty the most parsimonious approach.

James Andreoni
Department of Economics
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0508
and NBER
andreoni@ucsd.edu

Charles Sprenger
Department of Economics
Stanford University
Stanford, CA 94305
csprenge@ucsd.edu

1 Introduction

The theory of Expected Utility (EU) is among the most elegant and esthetically pleasing results in all of economics. It shows that if a preference ordering over a given set of gambles is complete, transitive, continuous, and, in addition, it satisfies the independence axiom, then utility is linear in objective probabilities.¹ The idea that a gamble's utility *could* be represented by the mathematical expectation of its utility outcomes dates to the St. Petersburg Paradox (Bernouilli, 1738). The idea that a gamble's utility was necessarily such an expectation if independence and the other axioms were satisfied became clear only in the 1950's (Samuelson, 1952, 1953).²

Two parallel research tracks have developed with respect to the independence axiom. The first takes linearity-in-probabilities as given and fits functional forms of the utility of consumption to experimental data (Birnbaum, 1992; Kachelmeier and Shehata, 1992; Holt and Laury, 2002).³ The second track focuses on identifying violations of independence.⁴ Principal among these are Allais' (1953b) well documented and extensively replicated common consequence and common ratio paradoxes.⁵

These and other violations of EU motivated new theoretical and experimental exercises, the most noteworthy being Cumulative Prospect Theory's (CPT) inverted *S*-

¹Subjective Expected Utility is not discussed in this paper. All results will pertain only to objective probabilities.

²The independence axiom is closely related to the Savage (1954) 'sure-thing principle' for subjective expected utility (Samuelson, 1952). Expected utility became known as von Neumann-Morgenstern (vNM) preferences after the publication of von Neumann and Morgenstern (1944). Independence, however, was not among the discussed axioms, but rather implicitly assumed. Samuelson (1952, 1953) discusses the resulting confusion and his suspicion of an implicit assumption of independence in the vNM treatment. Samuelson's suspicion was then confirmed in a note by Malinvaud (1952). For an excellent discussion of the history of the independence axiom, see Fishburn and Wakker (1995).

³Harrison and Rutstrom (2008) provide a detailed summary of both the experimental methods and estimation exercises associated with this literature.

⁴This second line of research began contemporaneously with the recognition of the importance of the independence axiom. Indeed Allais' presentation of Allais (1953a) was in the same session as Samuelson's presentation of Samuelson (1953) and the day after Savage's presentation of Savage (1953) at the Colloque Internationale d'Econométrie in Paris in May of 1952.

⁵In addition to the laboratory replications of Kahneman and Tversky (1979); Tversky and Kahneman (1992), there is now an extensive catalogue of Allais style violations of EU (Camerer, 1992; Harless and Camerer, 1994; Starmer, 2000).

shaped non-linear probability weighting (Kahneman and Tversky, 1979; Quiggin, 1982; Tversky and Kahneman, 1992; Tversky and Fox, 1995). In a series of experiments eliciting certainty equivalents for gambles, Tversky and Kahneman (1992) and Tversky and Fox (1995) estimated utility parameters supporting a model in which subjects “edit” probabilities by down-weighting high probabilities and up-weighting low probabilities. Identifying the S -shape of the weighting function and determining its parameter values has received significant attention both theoretically and in experiments (Wu and Gonzalez, 1996; Prelec, 1998; Gonzalez and Wu, 1999; Abdellaoui, 2000).⁶

Perhaps suprisingly, there are few direct tests of the independence axiom’s most critical implication: linearity-in-probabilities of the expected utility function. If, as in the first branch of literature, the independence axiom is assumed for identification of utility parameters, then the axiom cannot be rejected. Likewise, in the second branch, tests of probability weighting are not separate from functional form assumptions and thus are unlikely to confirm the independence axiom if it in fact holds unless, of course, both consumption utility and the weighting function are correctly specified.⁷ Other tests have, however, demonstrated important failures of EU beyond the Allais Paradox (Allais, 1953b). These include the calibration theorem (Rabin, 2000a,b), and goodness-of-fit comparisons (Camerer, 1992; Hey and Orme, 1994; Harless and Camerer, 1994).

In this paper, we provide a simple direct test of linearity-in-probabilities that reveals when independence holds, how it fails, and the nature of violations. We reintroduce an experimental method, which we call the *uncertainty equivalent*. Whereas a certainty equivalent identifies the certain amount that generates indifference to a given gamble, the uncertainty equivalent identifies the probability mixture over the gamble’s best outcome and zero that generates indifference. For example, consider a $(p, 1 - p)$ gamble

⁶Based upon the strength of these findings researchers have developed new methodology for eliciting risk preferences such as the ‘trade-off’ method (Wakker and Deneffe, 1996) that is robust to non-linear probability weighting.

⁷This observation is made by Abdellaoui (2000). Notable exceptions are the non-parametric probability weighting estimates of Gonzalez and Wu (1999); Bleichrodt and Pinto (2000) and Abdellaoui (2000) which find support for non-linearity-in-probabilities (see sub-section 4.2.1 for discussion).

over \$10 and \$30, $(p; 10, 30)$. The uncertainty equivalent identifies the $(q, 1 - q)$ gamble over \$30 and \$0, $(q; 30, 0)$, that generates indifference.⁸ Independence implies a linear relationship between p and q .

The uncertainty equivalent draws its motivation from the derivation of expected utility, where the cardinal index for a gamble is derived as the probability mixture over the best and worst options in the space of gambles.⁹ This means that in an uncertainty equivalent measure, the elicited q in $(q; Y, 0)$ can be interpreted as a utility index for the p gamble, $(p; X, Y)$, when $Y > X > 0$.

The uncertainty equivalent can also be used to inform the discussion of a variety of non-EU preference models including S -shaped probability weighting, expectations-based reference-dependence such as disappointment aversion (Bell, 1985; Loomes and Sugden, 1986; Gul, 1991; Koszegi and Rabin, 2006, 2007)¹⁰, and “ u - v ” preferences (Neilson, 1992; Schmidt, 1998; Diecidue, Schmidt and Wakker, 2004).¹¹ Though these models are often motivated by violations of independence and the Allais (1953b) paradox, they have both divergent psychological accounts of the phenomenon and divergent predictions in the uncertainty equivalent environment. Probability weighting explains

⁸We recognize that it is a slight abuse of traditional notation to have the probability refer to the lower outcome in the given gamble and the higher outcome in the uncertainty equivalent. It does, however, ease explication to have p refer to the probability of the low value and q refer to the probability of the high value.

⁹Such derivations are provided in most textbook treatments of expected utility. See, e.g. Varian (1992). Our research has uncovered that methods like our uncertainty equivalent were discussed in Farquhar’s (1984) excellent survey of utility assessment methods and, to our knowledge, were implemented experimentally in only one study of nine subjects using hypothetical monetary rewards (McCord and de Neufville, 1986), and a number of medical questionnaires (Magat, Viscusi and Huber, 1996; Oliver, 2005, 2007; Bleichrodt, Abellan-Perinan, Pinto-Prades and Mendez-Martinez, 2007).

¹⁰We include the Koszegi and Rabin (2006, 2007) model in the broad class of expectations-based reference dependence as the model’s predictions will closely resemble those of standard disappointment aversion in the present context as well as most other experimental environments (Ericson and Fuster, Forthcoming; Gill and Prowse, 2010; Abeler, Falk, Goette and Huffman, Forthcoming). For specific evidence distinguishing Koszegi and Rabin (2006, 2007) preferences from disappointment aversion, see Sprenger (2010).

¹¹ u - v preferences are less well-known than the other preference models. For a discussion of the early history of u - v preferences, see Schoemaker (1982). These models capture the intuition of Allais (1953b) that when options are far from certain, individuals act effectively as EU maximizers but, when certainty is available, it is disproportionately preferred. The u - v model differs in important ways from extreme or even discontinuous probability weighting and prior experiments have demonstrated these differences (Andreoni and Sprenger, 2011).

the Allais paradox with non-linear probability distortions, disappointment aversion relies instead on reference-dependence around an expectations-based reference point, and u - v preferences rely on a direct preference for certainty. Importantly, in the uncertainty equivalent environment these three models have different predictions that can be examined *without* relying on functional form assumptions for utility.

We conducted a within-subject experiment with 76 undergraduates at the University of California, San Diego, using both uncertainty and certainty equivalents. We find that that the model of Expected Utility does strikingly well, except near certainty. At probabilities of 0.95 and above, the independence axiom fails for about 40 percent of subjects, and in a very systematic way. A closer examination of the data strongly rejects prospect theory, and finds the data are most parsimoniously captured by a model of u - v preferences.

We demonstrate this finding in four steps. First, using uncertainty equivalents we find that p and q are, quite strikingly, related linearly for values of p away from certainty. Second, linearity breaks down as probabilities approach 1, and in a way that strongly contradicts S -shaped probability weighting. Third, 38 percent of subjects violate stochastic dominance at certainty, providing a within-subject example of the recently debated ‘uncertainty effect’ (Gneezy, List and Wu, 2006; Rydval, Ortmann, Prokoshcheva and Hertwig, 2009; Keren and Willemsen, 2008; Simonsohn, 2009). Such violations are a prediction of both the u - v model and some formulations of disappointment aversion, and are indicative of a disproportionate preference for certainty. Fourth, in the certainty equivalents tasks, subjects show both small stakes risk aversion and apparent S -shaped probability weighting, reproducing prior findings. We show, however, that these phenomena are driven by subjects who violate stochastic dominance in uncertainty equivalents. This suggests that extreme experimental risk aversion and probability weighting may be artifacts of a disproportionate preference for certainty coupled with specification error. For instance, a model of u - v preferences can pro-

duce certainty equivalent data that would generate an S -shaped weighting function if the econometrician estimated the functional forms of Tversky and coauthors, and also explain the data found in uncertainty equivalents. The opposite, however is not true.

Our experiment casts a new light on previous findings and raises new challenges for research on risk. The fact that we use non-parametric tests to find the independence axiom fails only near certainty, thus dramatically contradicting Cumulative Prospect Theory, shows the potential pitfalls of using parametric assumptions and model-fitting to identify a theory. The fact that a sizable minority of our subjects violate first order stochastic dominance at certainty is a second surprising but important element of the data. The violations are frequent and significant enough to deserve to be addressed by both further replication and by theoretical analysis. Both of these findings point to a special role for certainty in individual preferences.

Our experiments are, of course, not the first to recognize the importance of certainty. The original Allais (1953b) paradoxes drew attention to certainty being “disproportionately preferred,” and others have noted the preponderance of the evidence against EU also implicates certainty as an aider and abettor (Conlisk, 1989; Camerer, 1992; Harless and Camerer, 1994; Starmer, 2000). In related work on time preferences, Andreoni and Sprenger (2011), we document intertemporal results consistent with a specific preference for certainty. Recognizing that certainty may be disproportionately preferred gives a reason, perhaps, to expect non-EU behavior in experimental methodologies such as certainty equivalents, as certainty effects are built into the experimental design.¹²

The paper continues as follows. Section 2 discusses the uncertainty equivalent methodology and develops predictions based on different preference models. Section 3 presents experimental design details. Section 4 presents results and Section 5 concludes.

¹²Relatedly, one should also expect present-biased intertemporal behavior in designs that provide certain sooner payments and risky later payments. See Andreoni and Sprenger (2011) for further discussion.

2 The Uncertainty Equivalent

Consider a lottery $(p; X, Y)$ which provides $\$X$ with probability p and $\$Y > \X with probability $1 - p$. A certainty equivalent task elicits the certain amount, $\$C$, that is indifferent to this gamble. The uncertainty equivalent elicits the q -gamble over $\$Y$ and $\$0$, $(q; Y, 0)$, that is indifferent to this gamble. Take for example a 50%-50% gamble paying either $\$10$ or $\$30$. The uncertainty equivalent is the q -gamble over $\$30$ and $\$0$ that generates indifference.¹³

Under standard preference models, a more risk averse individual will, for a given gamble, have a lower certainty equivalent, C , and a higher uncertainty equivalent, q . A powerful distinction of the uncertainty equivalent is, however, that it is well-suited to identifying alternative preference models such as S -shaped probability weighting, where the non-linearity of the weighting function can be recovered directly, disappointment aversion, and u - v preferences. If preferences under certainty differ from those under uncertainty as in both disappointment aversion and u - v models, then certainty equivalent methodology assuming a single utility function is misspecified.¹⁴ Risk preferences or probability weights are not identified separately from differential preferences over certainty and uncertainty.

2.1 Empirical Predictions

We present empirical predictions in the uncertainty equivalent environment for expected utility, S -shaped probability weighting, disappointment aversion, and u - v preferences.¹⁵ Unlike experimental contexts that require functional form assumptions for

¹³We should be clear to distinguish the uncertainty equivalent from a probability equivalent. A probability equivalent elicits the probability of winning that makes a gamble indifferent to a *sure* amount. Uncertainty equivalents have risk on both sides of the indifference condition.

¹⁴In the u - v model, the differential preferences over certainty and uncertainty are delivered via a discontinuity. In disappointment aversion, choice under uncertainty involves disappointment and so an extra utility parameter not present in choice under certainty. See sub-sections 2.1.3 and 2.1.4 for further discussion.

¹⁵This is, of course, a limited list of the set of potentially testable decision models. For example, we do not discuss ambiguity aversion or the anticipatory utility specifications of Kreps and Porteus

model identification, the uncertainty equivalent can generally provide tests of utility specifications based on the relationship between p and q without appeal to specific functional form for utility.

2.1.1 Expected Utility

Consider a gamble with probability p of $\$X$ and probability $1 - p$ of a larger payment $\$Y > \X . The uncertainty equivalent of this prospect is the value q satisfying

$$p \cdot u(X) + (1 - p) \cdot u(Y) = q \cdot u(Y) + (1 - q) \cdot u(0).$$

Assuming $u(0) = 0$, $u(Y) > u(X)$, and letting $\theta = u(X)/u(Y) < 1$, then

$$q = p \cdot \frac{u(X)}{u(Y)} + 1 - p = 1 - p \cdot (1 - \theta),$$

and

$$\frac{dq}{dp} = \frac{u(X)}{u(Y)} - 1 = -(1 - \theta) < 0.$$

Thus, expected utility generates a negative *linear* relationship between the probability p of $\$X$ and the probability q of $\$Y$. This is an easily testable prediction.

2.1.2 Cumulative Prospect Theory Probability Weighting

Under Cumulative Prospect Theory, probabilities are weighted by the non-linear function $\pi(p)$. One popular functional form is the one parameter function used in Tversky and Kahneman (1992)¹⁶, $\pi(p) = p^\gamma / (p^\gamma + (1 - p)^\gamma)^{1/\gamma}$, $0 < \gamma < 1$. This inverted S-shaped function, as with others used in the literature, has the property that $\pi'(p)$ approaches infinity as p approaches 0 or 1. Probability weights are imposed on the

(1978) and Epstein and Zin (1989) as experimental uncertainty was resolved during the experimental sessions. These models generally reduce to expected utility when uncertainty is resolved immediately.

¹⁶Tversky and Fox (1995) and Gonzalez and Wu (1999) employ a similar two parameter $\pi(p)$ function. See Prelec (1998) for alternative specifications.

higher of the two utility values.¹⁷

Under this CPT formulation, the uncertainty equivalent indifference condition is

$$(1 - \pi(1 - p)) \cdot u(X) + \pi(1 - p) \cdot u(Y) = \pi(q) \cdot u(Y) + (1 - \pi(q)) \cdot u(0).$$

Again letting $u(0) = 0$ and $\theta = u(X)/u(Y) < 1$,

$$(1 - \pi(1 - p)) \cdot \theta + \pi(1 - p) = \pi(q).$$

This implicitly defines q as a function of p , yielding

$$\frac{dq}{dp} = -\frac{\pi'(1 - p)}{\pi'(q)} \cdot [1 - \theta] < 0.$$

As with expected utility, q and p are negatively related. Contrary to expected utility, the rate of change, dq/dp , depends on both p and q . Importantly, as p approaches 1, $\pi'(1 - p)$ approaches infinity and, provided finite $\pi'(q)$, the slope dq/dp becomes increasingly negative.¹⁸ This is a clearly testable alternative to expected utility. Importantly, the argument does not rest on the derivatives of the probability weighting function. Any modified S -shaped weighting function featuring up-weighting

¹⁷This formulation is assumed for binary gambles over strictly positive outcomes in Kahneman and Tversky (1979) and for all gambles in Tversky and Kahneman (1992). We abstract away from prospect theory's fixed reference point formulation as static reference points do not alter the analysis. Changing reference points, as in disappointment aversion, are discussed in sub-section 2.1.3.

¹⁸It is difficult to create a general statement for all probabilities between zero and one based upon the second derivative d^2q/dp^2 , as the second derivatives of the weighting function can be positive or negative depending on the concavity or convexity of the S -shaped distortion. The second derivative is

$$\frac{d^2q}{dp^2} = \frac{\pi''(1 - p) \cdot [1 - \theta] \cdot \pi'(q) + \pi''(q) \frac{dq}{dp} \cdot \pi'(1 - p) \cdot [1 - \theta]}{\pi'(q)^2}.$$

However, for p near 1, the sign is partially determined by $\pi''(q)$ which may be negative or positive. For S -shaped weighting, $\pi''(1 - p)$ will be negative in the concave region of low probabilities, and dq/dp will be negative from the development above. If q lies in the convex weighting region, such that $\pi''(q) > 0$, then $d^2q/dp^2 < 0$ and the relationship is concave and may remain so with $\pi''(q) < 0$. Consensus puts the concave region between probability 0 and around 1/3 (Tversky and Kahneman, 1992; Tversky and Fox, 1995; Prelec, 1998). As will be seen, the uncertainty equivalents for $p = 1$ lie substantially above 1/3 for all of our experimental conditions such that a concave relationship between p and q would be expected.

of low probabilities and down-weighting of high probabilities will share the characteristic that the relationship between q and p will become more negative as p approaches 1.¹⁹

2.1.3 Disappointment Aversion

We use the term disappointment aversion to refer to a broad class of reference-dependent models with expectations-based reference points where a gamble's outcomes are evaluated relative to the gamble's EU certainty equivalent (Bell, 1985; Loomes and Sugden, 1986; Gul, 1991). Recent research on expectations-based reference dependence extends the notion of reference points to reference distributions (Koszegi and Rabin, 2006, 2007). In the environment described in this paper, models with reference distributions and models with reference points generate very similar predictions. For simplicity, we present the analysis in terms of expectations-based reference points.²⁰

Consider a p -gamble over $\$X$ and $\$Y > \X . Let C_p be the standard EU certainty equivalent satisfying

$$p \cdot u(X) + (1 - p) \cdot u(Y) = u(C_p).$$

Taking C_p as the reference point, the reference-dependent utility of the p -gamble is

$$U_p = p \cdot \tilde{u}(X|C_p) + (1 - p) \cdot \tilde{u}(Y|C_p)$$

where $\tilde{u}(\cdot|C_p)$ is the reference-dependent utility function with a reference point at C_p .

We assume a standard specification for $u(\cdot|C_p)$ (Bell, 1985; Loomes and Sugden,

¹⁹This would be the case for virtually all functional forms and parameter values discussed in Prelec (1998) and for functions respecting condition (A) of the Quiggin (1982) weighting function. Take p close to 1 and $(1 - p)$ close to zero, $u(Y)$ will be up-weighted and $u(X)$ will be down-weighted on the left hand side of the above indifference condition. In order to compensate for the up-weighting of the good outcome on the left hand side, q on the right hand side must be high. At $p = 1$, the up-weighting of $u(Y)$ disappears precipitously and so q decreases precipitously to maintain indifference.

²⁰For analysis focusing on the distinction between Koszegi and Rabin (2006, 2007) preferences and other models of disappointment aversion as well as discussion of the different equilibrium and utility maximization concepts across the models, see Sprenger (2010).

1986),

$$\tilde{u}(z|C_p) = u(z) + \mu(u(z) - u(C_p)),$$

where the function $u(z)$ represents consumption utility for some outcome, z , and $\mu(\cdot)$ represents disappointment-elation utility relative to the referent, C_p . Several simplifying assumptions are made. Following Koszegi and Rabin (2006, 2007) we assume a piecewise-linear disappointment-elation function,

$$\mu(u(z) - u(C_p)) = \left\{ \begin{array}{ll} \eta \cdot (u(z) - u(C_p)) & \text{if } u(z) - u(C_p) \geq 0 \\ \eta \cdot \lambda \cdot (u(z) - u(C_p)) & \text{if } u(z) - u(C_p) < 0 \end{array} \right\},$$

where the utility parameter with $\lambda > 1$ indicates disappointment aversion. Without loss of generality, let $\eta = 1$. The disappointment averse utility of the $(p; X, Y)$ gamble can be written as

$$U_p = p \cdot [u(X) + \lambda \cdot (u(X) - u(C_p))] + (1 - p) \cdot [u(Y) + 1 \cdot (u(Y) - u(C_p))].$$

Replacing $u(C_p) = p \cdot u(X) + (1 - p) \cdot u(Y)$, this becomes

$$U_p = [p + p \cdot (1 - p) \cdot (\lambda - 1)] \cdot u(X) + [(1 - p) - p \cdot (1 - p) \cdot (\lambda - 1)] \cdot u(Y).$$

Note that this implies that disappointment aversion is another version of a probability weighted utility function with weighting function

$$\tilde{\pi}(1 - p) = (1 - p) - p \cdot (1 - p) \cdot (\lambda - 1).$$

In addition $\tilde{\pi}(1 - p) \leq 1 - p$ if $\lambda > 1$, and $\tilde{\pi}(1 - p)$ is a convex function.²¹ Hence disappointment aversion is equivalent to a specific form of probability weighting that

²¹ $\tilde{\pi}(1 - p)$ describes a parabola with a critical point at $1 - p = (\lambda - 2)/(2\lambda - 2)$.

is not S -shaped, but rather downweights all probabilities.²² Following identical logic to that of the previous section the uncertainty equivalent indifference relation again implies

$$\frac{dq}{dp} = -\frac{\tilde{\pi}'(1-p)}{\tilde{\pi}'(q)} \cdot [1-\theta].$$

Because the weighting function, $\tilde{\pi}(\cdot)$, is convex, one can easily check that the second derivative, $\frac{d^2q}{dp^2}$, is greater than zero, implying that disappointment aversion predicts a convex relationship between p and q for $\lambda > 1$.²³

Of additional interest is that as p approaches 1, $\tilde{\pi}'(1-p)$ approaches $2-\lambda$ under our formulation. For sufficiently disappointment averse individuals ($\lambda > 2$ in this example) the relationship between p and q will become positive as p approaches 1, provided $\tilde{\pi}'(q) > 0$. This is an important prediction of disappointment aversion. A positive relationship between p and q near certainty implies violations of first order stochastic

²²Loomes and Sugden (1986) and Gul (1991) provide similar demonstrations that disappointment aversion is observationally equivalent to down-weighting of all probabilities.

²³Note that $\tilde{\pi}(q) \geq \tilde{\pi}(1-p)$ is implied from the above indifference condition, and, for a weakly increasing $\tilde{\pi}(\cdot)$, $q \geq 1-p$. Convexity implies $\tilde{\pi}'(q) \geq \tilde{\pi}'(1-p)$. For the employed specification $\tilde{\pi}'(1-p) = 2-\lambda + 2(1-p)(\lambda-1)$ and $\tilde{\pi}''(\cdot)$ is a constant, such that $\tilde{\pi}''(1-p) = \tilde{\pi}''(q) = 2(\lambda-1)$. This second derivative is positive under the assumption $\lambda > 1$. Hence, the sign of

$$\frac{d^2q}{dp^2} = \frac{\tilde{\pi}''(1-p) \cdot [1-\theta] \cdot \tilde{\pi}'(q) + \tilde{\pi}''(q) \frac{dq}{dp} \cdot \tilde{\pi}'(1-p) \cdot [1-\theta]}{\tilde{\pi}'(q)^2}.$$

depends on the sign of

$$\tilde{\pi}'(q) + \frac{dq}{dp} \cdot \tilde{\pi}'(1-p).$$

Plugging in for dq/dp

$$\tilde{\pi}'(q) - \frac{\tilde{\pi}'(1-p)}{\tilde{\pi}'(q)} \cdot [1-\theta] \cdot \tilde{\pi}'(1-p),$$

and dividing by $\tilde{\pi}'(q)$ we obtain

$$1 - \frac{\tilde{\pi}'(1-p)}{\tilde{\pi}'(q)} \cdot [1-\theta] \cdot \frac{\tilde{\pi}'(1-p)}{\tilde{\pi}'(q)}.$$

Because convexity of $\tilde{\pi}(\cdot)$ and $q \geq (1-p)$ implies $\tilde{\pi}'(q) \geq \tilde{\pi}'(1-p)$, $\tilde{\pi}'(1-p)/\tilde{\pi}'(q) \leq 1$. Additionally $1-\theta < 1$, by the assumption of monotonicity. The second term is therefore a multiplication of three terms that are less than or equal to 1 and one concludes

$$1 - \frac{\tilde{\pi}'(1-p)}{\tilde{\pi}'(q)} \cdot [1-\theta] \cdot \frac{\tilde{\pi}'(1-p)}{\tilde{\pi}'(q)} > 0,$$

$d^2q/dp^2 > 0$, the relationship is convex.

dominance as certainty is approached. The uncertainty equivalent, q , acts as a utility index of the p -gamble. Given two offered gambles, p and p' , with $p > p'$, and two associated uncertainty equivalents, q and q' , a subject violates first order stochastic dominance if $q > q'$ as this indirectly reveals a preference for a gamble with higher probability of a lower prize.

Importantly, some disappointment averse models are constructed with assumptions guaranteeing that the underlying preferences satisfy stochastic dominance, taking violations of stochastic dominance as a disqualifying feature of a model of behavior (Loomes and Sugden, 1986; Gul, 1991), while others are not (Bell, 1985; Koszegi and Rabin, 2006, 2007). Hence, whether the data show significant violations of stochastic dominance can be informative for the constraints placed on models of expectations-based reference dependence.

2.1.4 u - v Preferences

The u - v model (Neilson, 1992; Schmidt, 1998; Diecidue et al., 2004) is designed to capture Allais' (1953b) intuition of a disproportionate preference for security in the 'neighborhood of certainty.' Let $u(X)$ be the utility of $\$X$ with uncertainty and $v(X)$ be the utility of $\$X$ with certainty. Assume $v(X) > u(X)$ for $\$X > 0$. Under such u - v preferences, p and q will have a linear relationship away from $p = 1$. At $p = 1$, the discontinuity in utility introduces a discontinuity in the relationship between p and q . At $p = 1$, the q that solves the indifference condition

$$v(X) = q \cdot u(Y)$$

will be

$$q = \frac{v(X)}{u(Y)} > \frac{u(X)}{u(Y)}.$$

With the u - v specification, q will be linearly decreasing in p and then discontinuously *increase* at $p = 1$.

Importantly, if the neighborhood of certainty is understood to begin at probabilities less than one, the discontinuity may appear below $p = 1$. Instead of a discontinuity at certainty, however, one could imagine a less rigid model of preferences where the relationship between p and q is continuous but non-linear as p approaches 1. Distinguishing between such a representation and disappointment aversion would be virtually impossible. In this sense, discontinuous u - v preferences could act as a simple, parsimonious representation of disappointment averse decision-making without making appeals to an expectations-based reference point. All uncertainty entails disappointment and so lower utility. This is similar in spirit to the β - δ representation of time preferences to usefully approximate hyperbolic discounting.

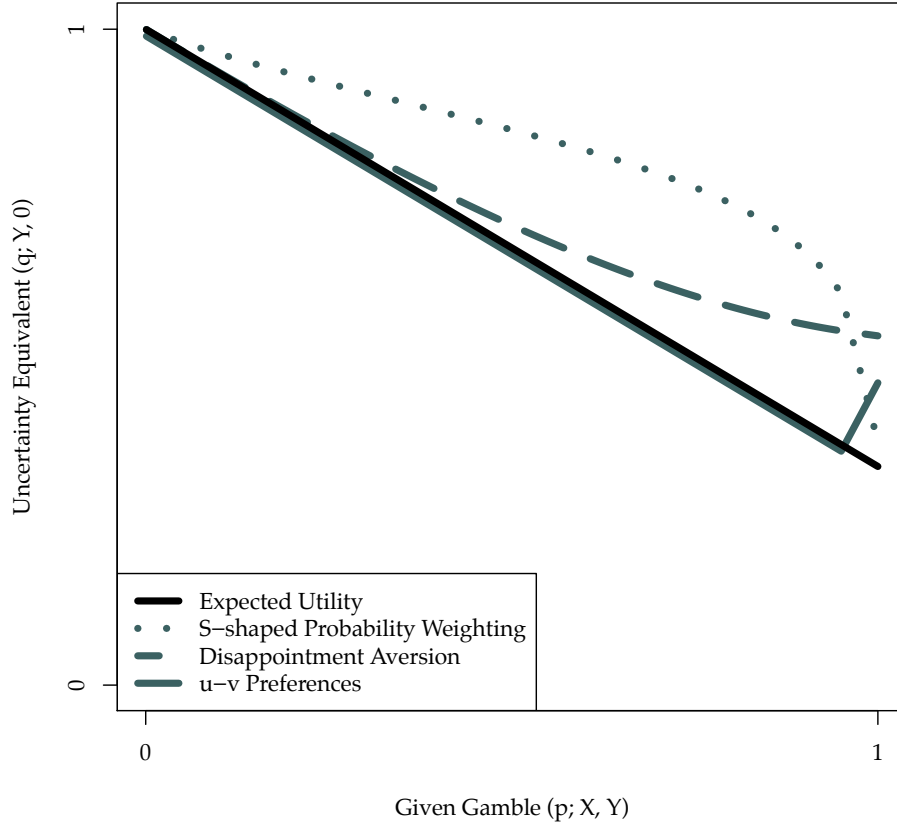
Similar to some versions of disappointment aversion, the u - v preference model violates first order stochastic dominance if there exists a disproportionate preference for certainty, $v(X) > u(X)$. Certainty of a small payment will be preferred to near certain gambles paying this small amount with sufficiently high probability and something larger with low probability.

Though such a property can be viewed as a weakness of the u - v preference model (Diecidue et al., 2004), we again take the view that violating stochastic dominance is an implication of the model that can easily be used to test and reject it.

2.2 Summary

Figure 1 presents the theoretical predictions of the four models just discussed. Importantly, the uncertainty equivalent environment provides clear separation between the models. Under expected utility, q should be a linear function of p . Under S -shaped probability weighting q should be a concave function of p with the relationship growing more negative as p approaches 1. Under disappointment aversion q should be a convex

Figure 1: Empirical Predictions



Note: Empirical predictions of the relationship between given gambles, $(p; X, Y)$, and uncertainty equivalents $(q; Y, 0)$ for Expected Utility, S -shaped CPT probability weighting, disappointment aversion, and u - v preferences. A linear prediction is obtained for EU, a concave relationship for S -shaped CPT probability weighting, and a convex relationship for disappointment aversion. For u - v preferences a linear negative relationship between $(p; X, Y)$ and $(q; Y, 0)$ is obtained for $p < 1$, with a discontinuous increase in $(q; Y, 0)$ at certainty, $p = 1$.

function of p , perhaps with sharper convexity as p approaches 1, and with indirect violations of stochastic dominance. Under u - v preferences, q should be a linear function of p until certainty, with convexity appearing near certainty and being associated with indirect violations of stochastic dominance.

3 Experimental Design

Eight uncertainty equivalents were implemented with probabilities $p \in \{0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95, 1\}$ in three different payment sets, $(X, Y) \in \{(10, 30), (30, 50), (10, 50)\}$, yielding 24 total uncertainty equivalents. The experiment was conducted with paper-and-pencil and each payment set (X, Y) was presented as a packet of 8 pages. The uncertainty equivalents were presented in increasing order from $p = 0.05$ to $p = 1$ in a single packet.

On each page, subjects were informed that they would be making a series of decisions between two options. Option A was a p chance of receiving $\$X$ and a $1 - p$ chance of receiving $\$Y$. Option A remained the same throughout the page. Option B varied in steps from a 5 percent chance of receiving $\$Y$ and a 95 percent chance of receiving $\$0$ to a 99 percent chance of receiving $\$Y$ and a 1 percent chance of receiving $\$0$. Figure 2 provides a sample decision task. In this price list style experiment, the row at which a subject switches from preferring Option A to Option B indicates the range of values within which the uncertainty equivalent, q , lies.

A common frustration with price lists is that anywhere from 10 to 50 percent of subjects can be expected to switch columns multiple times.²⁴ Because such multiple switch points are difficult to rationalize and may indicate subject confusion, it is common for researchers to drop these subjects from the sample.²⁵ Instead, we augmented the standard price list with a simple framing device designed to clarify the decision process. In particular, we added a line to both the top and bottom of each price list in which the choices were clear, and illustrated this by checking the obvious best option. The top line shows that each p -gamble is preferred to a 100 percent chance of receiving $\$0$ while the bottom line shows that a 100 percent chance of receiving $\$Y$ is

²⁴Holt and Laury (2002) and had around 10 percent and Jacobson and Petrie (2009) had nearly 50 percent multiple switchers. An approximation of a typical fraction of subjects lost to multiple switch points in an MPL is around 15 percent.

²⁵Other options include selecting one switch point to be the “true point” (Meier and Sprenger, 2010) or constraining subjects to a single switch point (Harrison, Lau, Rutstrom and Williams, 2005).

Figure 2: Sample Uncertainty Equivalent Task

	Option A		<i>or</i>	Option B			
	Chance of \$10	Chance of \$30		Chance of \$0	Chance of \$30		
	50 in 100	50 in 100	<input checked="" type="checkbox"/>	<i>or</i>	100 in 100	0 in 100	<input type="checkbox"/>
1)	50 in 100	50 in 100	<input type="checkbox"/>	<i>or</i>	95 in 100	5 in 100	<input type="checkbox"/>
2)	50 in 100	50 in 100	<input type="checkbox"/>	<i>or</i>	90 in 100	10 in 100	<input type="checkbox"/>
3)	50 in 100	50 in 100	<input type="checkbox"/>	<i>or</i>	85 in 100	15 in 100	<input type="checkbox"/>
4)	50 in 100	50 in 100	<input type="checkbox"/>	<i>or</i>	80 in 100	20 in 100	<input type="checkbox"/>
5)	50 in 100	50 in 100	<input type="checkbox"/>	<i>or</i>	75 in 100	25 in 100	<input type="checkbox"/>
6)	50 in 100	50 in 100	<input type="checkbox"/>	<i>or</i>	70 in 100	30 in 100	<input type="checkbox"/>
7)	50 in 100	50 in 100	<input type="checkbox"/>	<i>or</i>	65 in 100	35 in 100	<input type="checkbox"/>
8)	50 in 100	50 in 100	<input type="checkbox"/>	<i>or</i>	60 in 100	40 in 100	<input type="checkbox"/>
9)	50 in 100	50 in 100	<input type="checkbox"/>	<i>or</i>	55 in 100	45 in 100	<input type="checkbox"/>
10)	50 in 100	50 in 100	<input type="checkbox"/>	<i>or</i>	50 in 100	50 in 100	<input type="checkbox"/>
11)	50 in 100	50 in 100	<input type="checkbox"/>	<i>or</i>	45 in 100	55 in 100	<input type="checkbox"/>
12)	50 in 100	50 in 100	<input type="checkbox"/>	<i>or</i>	40 in 100	60 in 100	<input type="checkbox"/>
13)	50 in 100	50 in 100	<input type="checkbox"/>	<i>or</i>	35 in 100	65 in 100	<input type="checkbox"/>
14)	50 in 100	50 in 100	<input type="checkbox"/>	<i>or</i>	30 in 100	70 in 100	<input type="checkbox"/>
15)	50 in 100	50 in 100	<input type="checkbox"/>	<i>or</i>	25 in 100	75 in 100	<input type="checkbox"/>
16)	50 in 100	50 in 100	<input type="checkbox"/>	<i>or</i>	20 in 100	80 in 100	<input type="checkbox"/>
17)	50 in 100	50 in 100	<input type="checkbox"/>	<i>or</i>	15 in 100	85 in 100	<input type="checkbox"/>
18)	50 in 100	50 in 100	<input type="checkbox"/>	<i>or</i>	10 in 100	90 in 100	<input type="checkbox"/>
19)	50 in 100	50 in 100	<input type="checkbox"/>	<i>or</i>	5 in 100	95 in 100	<input type="checkbox"/>
20)	50 in 100	50 in 100	<input type="checkbox"/>	<i>or</i>	1 in 100	99 in 100	<input type="checkbox"/>
	50 in 100	50 in 100	<input type="checkbox"/>	<i>or</i>	0 in 100	100 in 100	<input checked="" type="checkbox"/>

Note: Sample uncertainty equivalent task for $(p; X, Y) = (0.5, 10, 30)$ eliciting $(q; 30, 0)$.

preferred to each p -gamble. These pre-checked gambles were not available for payment, but were used to clarify the decision task. This methodology is close to the clarifying instructions from the original Holt and Laury (2002), where subjects were described a 10 lottery choice task and random die roll payment mechanism and then told, “In fact, for Decision 10 in the bottom row, the die will not be needed since each option pays the highest payoff for sure, so your choice here is between 200 pennies or 385 pennies.”

Since the economist is primarily interested in the price list method as a means of measuring a single choice – the switching point – it seemed natural to include language

to this end. Hence, in directions subjects were told “Most people begin by preferring Option A and then switch to Option B, so one way to view this task is to determine the best row to switch from Option A to Option B.” Our efforts appear to have reduced the volume of multiple switching dramatically, to less than 1 percent of total responses. Observations with multiple switch points were removed from analysis and are noted.

In order to provide an incentive for truthful revelation of uncertainty equivalents, subjects were randomly paid one of their choices in cash at the end of the experimental session.²⁶ This random-lottery mechanism, which is widely used in experimental economics, does introduce a compound lottery to the decision environment. Starmer and Sugden (1991) demonstrate that this mechanism does not create a bias in experimental response. Seventy-six subjects were recruited from the undergraduate population at University of California, San Diego. The experiment lasted about one hour and average earnings were \$24.50, including a \$5 minimum payment.

3.1 Additional Risk Preference Measures

In addition to the uncertainty equivalents discussed above, subjects were also administered two Holt and Laury (2002) risk measures over payment values of \$10 and \$30 as well as 7 standard certainty equivalents tasks with p gambles over \$30 from the set $p \in \{0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95, 1\}$. These probabilities are identical to those used in the original probability weighting experiments of Tversky and Kahneman (1992) and Tversky and Fox (1995). The certainty equivalents were also presented in price list style with similar language to the uncertainty equivalents and could also be chosen for payment.²⁷ Examples of these additional risk measures are provided in the appendix. Two orders of the tasks were implemented: 1) UE, HL, CE and 2) CE,

²⁶Please see the instructions in the Appendix for payment information provided to subjects.

²⁷Multiple switching was again greatly reduced relative to prior studies to less than 1 percent of responses. Observations with multiple switch points were removed from analysis and are noted. As will be seen, results of the CE task reproduce the results of others. This increases our confidence that our innovations with respect to the price lists did not result in biased or peculiar measurement of behavior.

HL, UE to examine order effects, and none were found.²⁸

4 Results

We present our analysis in three sub-sections. First, we look at the uncertainty equivalents and provide tests of four models of utility. Second, we consider the standard certainty equivalents, reproducing the usual probability weighting phenomenon, and contradicting the results of the first sub-section. Third, we reconcile the uncertainty and certainty equivalent data, showing that a parsimonious model of a special preference for certainty can explain both sets of results, while Prospect Theory cannot.

4.1 Uncertainty Equivalents and Tests of Linearity

To provide estimates of the mean uncertainty equivalent and the appropriate standard error for each of the 24 uncertainty equivalent tasks, we first estimate non-parametric interval regressions (Stewart, 1983).²⁹ The interval response of q is regressed on indicators for all probability and payment-set interactions with standard errors clustered on the subject level. We calculate the relevant coefficients as linear combinations of interaction terms and present these in Table 1, Panel A. Figure 3 graphs the corresponding mean uncertainty equivalent, q , for each p , shown as dots with error bars.³⁰

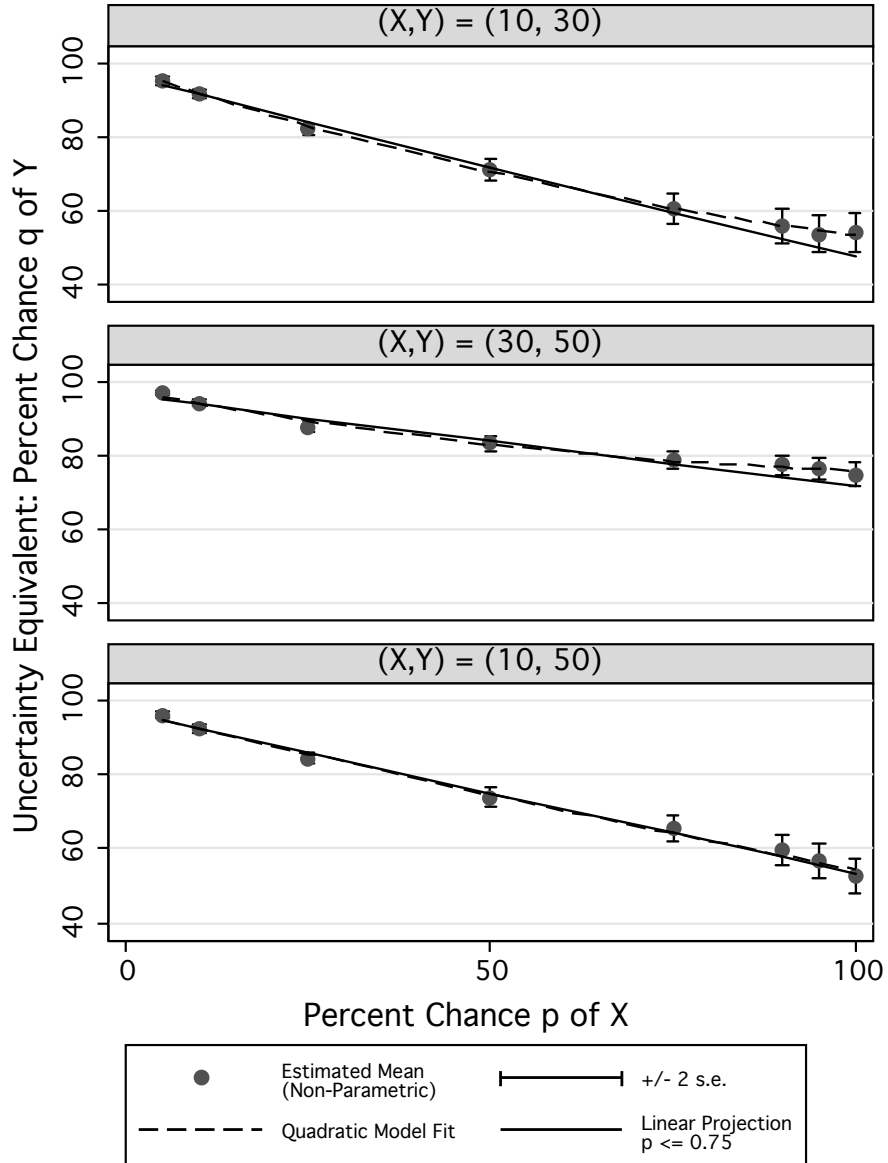
The first question we ask is: are p and q in an exact linear relationship, as predicted

²⁸Though we used the HL task primarily as a buffer between certainty and uncertainty equivalents, a high degree of correlation is obtained across elicitation techniques. As the paper is already long, correlations with HL data are discussed primarily in footnotes.

²⁹Identical results are obtained when using OLS and the midpoint of the interval.

³⁰Uncertainty equivalents correlate significantly with the number of safe choices chosen in the Holt-Laury risk tasks. For example, for $p = 0.5$ the individual correlations between the uncertainty equivalent q and the number of safe choices, S_{10} , in the \$10 HL task are $\rho_{q(10,30),S_{10}} = 0.52$ ($p < 0.01$), $\rho_{q(30,50),S_{10}} = 0.38$ ($p < 0.01$), and $\rho_{q(10,50),S_{10}} = 0.54$ ($p < 0.01$). The individual correlations between the uncertainty equivalent, q , and the number of safe choices, S_{30} , in the \$30 HL task are $\rho_{q(10,30),S_{30}} = 0.54$ ($p < 0.01$), $\rho_{q(30,50),S_{30}} = 0.45$ ($p < 0.01$), and $\rho_{q(10,50),S_{30}} = 0.67$ ($p < 0.01$). The correlation between the number of safe choices in the HL tasks is also high, $\rho_{S_{10},S_{30}} = 0.72$ ($p < 0.01$). These results demonstrate consistency across elicitation techniques as higher elicited q and a higher number of safe HL choices both indicate more risk aversion.

Figure 3: Uncertainty Equivalent Responses



Note: Figure presents uncertainty equivalent, $(q; Y, 0)$, corresponding to Table 1, Panel A for each given gamble, $(p; X, Y)$, of the experiment. The dashed black line represents the quadratic model fit of Table 1, Panel B. The solid black line corresponds to a linear projection based upon data from $p \leq 0.75$, indicating the degree to which the data adhere to the expected utility prediction of linearity away from certainty.

by expected utility? To answer this we conducted a linear interval regression of q on p for only those $p \leq 0.75$, with a linear projection to $p = 1$. This is presented as the solid

line in Figure 3. Figure 3 shows a clear pattern. The data fit the linear relationship extremely well for the bottom panel, the $(X, Y) = (10, 50)$ condition, but as we move up the linear fit begins to fail for probabilities of 0.90 and above, and becomes increasingly bad as p approaches certainty. In the $(10, 30)$ condition (top panel), EU fails to the point that the mean behavior violates stochastic dominance: the q for $p = 1$ is above the q for $p = 0.95$. Since q is a utility index for the p -gamble, this implies that a low outcome of \$10 for sure is worth more than a gamble with a 95 percent chance of \$10 and a 5 percent chance of \$30.

To explore the apparent non-linearity near $p = 1$, Table 1, Panels B and C present estimates of the relationship between q and p . Panel B estimates interval regressions assuming a quadratic relationship, and Panel C assumes a linear relationship. Expected utility is consistent with a square term of zero, S -shaped probability weighting with a negative square term, and disappointment aversion and u - v preferences with a positive square term. Panel B reveals a zero square term for the $(10, 50)$ condition, but positive and significant square terms for both $(30, 50)$ and $(10, 30)$ conditions.³¹

The parametric specifications of Panels B and C can be compared to the non-parametric specification presented in Panel A with simple likelihood ratio chi-square tests. Neither the quadratic nor the linear specification can be rejected relative to the fully non-parametric model: $\chi^2(15)_{A,B} = 8.23$, ($p = 0.91$); $\chi^2(18)_{A,C} = 23.66$, ($p = 0.17$). However, the linear specification of Panel C can be rejected relative to the parsimonious quadratic specification of Panel B, $\chi^2(3)_{B,C} = 15.43$, ($p < 0.01$). We reject expected utility's linear prediction in favor of a convex relationship between p and q .

Our results are important for evaluating linearity-in-probabilities, and for under-

³¹One can also interpret the coefficients on $p \times 100$ as a measure of utility function curvature at $p = 0$ where $dq/dp = -(1 - u(X)/u(Y))$. Under risk neutrality, this coefficient should be -0.66 for $(X, Y) = (10, 30)$, -0.4 for $(X, Y) = (30, 50)$ and -0.8 for $(X, Y) = (10, 50)$. Though the estimates in the $(X, Y) = (10, 30)$ and $(30, 50)$ conditions are close to the risk neutral prediction, the $(X, Y) = (10, 50)$ condition differs substantially from risk neutrality.

Table 1: Estimates of the Relationship Between q and p

	(1)	(2)	(3)
	$(X, Y) = (\$10, \$30)$	$(X, Y) = (\$30, \$50)$	$(X, Y) = (\$10, \$50)$
<i>Dependent Variable: Interval Response of Uncertainty Equivalent ($q \times 100$)</i>			
Panel A: Non-Parametric Estimates			
$p \times 100 = 10$	-3.623*** (0.291)	-2.575*** (0.321)	-3.869*** (0.413)
$p \times 100 = 25$	-13.270*** (0.719)	-8.867*** (0.716)	-11.840*** (0.748)
$p \times 100 = 50$	-24.119*** (1.476)	-13.486*** (0.916)	-22.282*** (1.293)
$p \times 100 = 75$	-34.575*** (2.109)	-17.790*** (1.226)	-30.769*** (1.777)
$p \times 100 = 90$	-39.316*** (2.445)	-19.171*** (1.305)	-36.463*** (2.190)
$p \times 100 = 95$	-41.491*** (2.635)	-20.164*** (1.411)	-39.721*** (2.425)
$p \times 100 = 100$	-41.219*** (2.626)	-21.747*** (1.536)	-43.800*** (2.454)
Constant	95.298*** (0.628)	96.822*** (0.290)	96.230*** (0.497)
<i>Log-Likelihood = -4498.66</i>			
<i>AIC = -9047.32, BIC = 9185.02</i>			
Panel B: Quadratic Estimates			
$p \times 100$	-0.660*** (0.060)	-0.376*** (0.035)	-0.482*** (0.047)
$(p \times 100)^2$	0.002*** (0.001)	0.002*** (0.000)	0.001 (0.000)
Constant	98.125*** (0.885)	97.855*** (0.436)	97.440*** (0.642)
<i>Log-Likelihood = -4502.77</i>			
<i>AIC = -9025.55, BIC = 9080.63</i>			
Panel C: Linear Estimates			
$p \times 100$	-0.435*** (0.027)	-0.209*** (0.016)	-0.428*** (0.027)
Constant	95.091*** (0.678)	95.603*** (0.512)	96.718*** (0.714)
<i>Log-Likelihood = -4510.49</i>			
<i>AIC = -9034.98, BIC = 9073.54</i>			

Notes: Coefficients from single interval regression for each panel (Stewart, 1983) with 1823 observations. Standard errors clustered at the subject level in parentheses. 76 clusters. The regressions feature 1823 observations because one individual had a multiple switch point in one uncertainty equivalent in the $(X, Y) = (\$10, \$50)$ condition.

Level of significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

standing the robustness of the standard probability weighting phenomenon. The data indicate that expected utility performs well away from certainty where the data adhere closely to linearity. However, the data deviate from linearity as p approaches 1, generating a convex relationship between p and q . This is a strong and significant rejection of the S -shaped probability weighting model. The finding is notable as the uncertainty equivalent is only a small deviation from standard certainty equivalents, where probability weighting has often been demonstrated.

While the data reject S -shaped probability weighting, both disappointment aversion and u - v preferences predict the convex relationship between p and q with sharpened convexity at $p = 1$. The difference between the models arises in that u - v preferences predicts a strictly linear relationship away from certainty while disappointment aversion predicts convexity throughout. Though the data adhere closely to linearity for $p \leq 0.75$ in Figure 3, significant positive square terms are obtained for $p \leq 0.75$ in regressions corresponding to Table 1, and the linear specification can be rejected relative to the quadratic specification, $\chi^2(3) = 20.07$, ($p < 0.01$). Supporting disappointment aversion, we reject linearity for probabilities away from certainty. However, linearity does provide surprisingly good model fit. Hence, EU can parsimoniously explain the data in this region, suggesting that deviations from EU, though econometrically significant, may not be economically significant away from certainty.

The analysis of this sub-section generates two results. First, expected utility performs remarkably well away from certainty. Second, at certainty behavior deviates from expected utility in a surprising way, rejecting Cumulative Prospect Theory.³² Taken

³²Hints of these results exist in the prior literature. McCord and de Neufville (1986), with nine experimental subjects and a related construct they called a lottery equivalent, document no systematic difference in utilities elicited below probability 1, but that elicited utility at probability one was “consistently above and to the right of the other functions” [p. 60]. Bleichrodt et al. (2007) use five methods of utility elicitation for health outcomes including certainty equivalents and lottery equivalents. Expected utility was found to perform poorly in decisions involving certainty, but well in comparisons involving only uncertain prospects. Additionally the utilities elicited with certainty were generally above those elicited with uncertainty. Though these results and other certainty effects are often argued to be supportive of S -shaped probability weighting, careful consideration of our results suggests otherwise.

together, the data are most consistent with disappointment aversion, though $u-v$ could also provide a parsimonious explanation. The challenge for some formulations of disappointment aversion is the significant presence of violations of first order stochastic dominance as p approaches 1. Next, we explore this in detail.

4.1.1 Violations of Stochastic Dominance

A substantial portion of our subjects violate first order stochastic dominance. These violations are organized close to certainty consistent with $u-v$ preferences and some formulations of disappointment aversion. Since the q elicited in an uncertainty equivalent acts as a utility index, dominance violations are identified when a subject reports a higher q for a higher p , indicating that they prefer a greater chance of a *smaller* prize.

Each individual has 84 opportunities to violate first order stochastic dominance in such a way.³³ We can identify the percentage of choices violating stochastic dominance at the individual level and so develop an individual violation rate. To begin, away from certainty, violations of stochastic dominance are few, averaging only 4.3% (*s.d.* = 6.4%). In the 21 cases per subject when certainty, $p = 1$, is involved, the individual violation rate increases significantly to 9.7% (15.8%), ($t = 3.88$, $p < 0.001$). When examining only the three comparisons of $p = 1$ to $p' = 0.95$, the individual violation rate increases further to 17.5% (25.8%), ($t = 3.95$, $p < 0.001$). Additionally, 38 percent (29 of 76) of subjects demonstrate at least one violation of stochastic dominance when comparing $p = 1$ to $p' = 0.95$. This finding suggests that violations of stochastic dominance are prevalent and tend to be localized close to certainty.³⁴

³³Identifying violations in this way recognizes the interval nature of the data as it is determined by price list switching points. We consider violations within each payment set (X, Y) . With 8 probabilities in each set, seven comparison can be made for $p = 1 : p' \in \{0.95, 0.9, 0.75, 0.5, 0.25, 0.1, 0.05\}$. Six comparisons can be made for $p = 0.95$ and so on, leading to 28 comparisons for each payment set and 84 within-set comparisons of this form.

³⁴It is important to note that the violations of stochastic dominance that we document are indirect measures of violation. We hypothesize that violations of stochastic dominance would be less prevalent in direct preference rankings of gambles with a dominance relation. Though we believe the presence of dominance violations can be influenced by frames, this is likely true for the presence of many decision phenomena. In the following sub-section we present data from certainty equivalents

To simplify discussion, we will refer to individuals who violate stochastic dominance between $p = 1$ and $p' = 0.95$ as *Certainty Preferring*. The remaining 62 percent of subjects are classified as *Certainty Neutral*.³⁵

Figure 4 reproduces Figure 3, but splits the sample by certainty preference. First, this shows the roughly 60% of subjects that are classified as Certainty Neutral demonstrate a linear relationship between q and p throughout. In estimates corresponding to Table 1, Panel B, negligible and insignificant square terms are obtained and quadratic and linear specifications cannot be distinguished ($\chi^2(3)_{B,C} = 0.69$, $p = 0.88$).³⁶ These data show that without a specific minority of individuals who exhibit a disproportionate preference for certainty, expected utility organizes the data extremely well. This finding of linearity is additionally important because eliminating the convexity of Certainty Preferring individuals should, in principle, give S -shaped probability weighting's concave prediction the best opportunity to be revealed.

Second, the mean uncertainty equivalents in Figure 4, Panels A and B coincide away from certainty and decline linearly with p . However, the uncertainty equivalents for subjects with a disproportionate preference for certainty peel away as certainty is approached.

Third, for Certainty Preferring subjects aggregate violations of stochastic dominance are less pronounced in the $(X, Y) = (10, 50)$ condition. Andreoni and Sprenger (2009c) discuss experimental conditions when violations of stochastic dominance are more or less likely to be observed in experimental data and demonstrate that for one

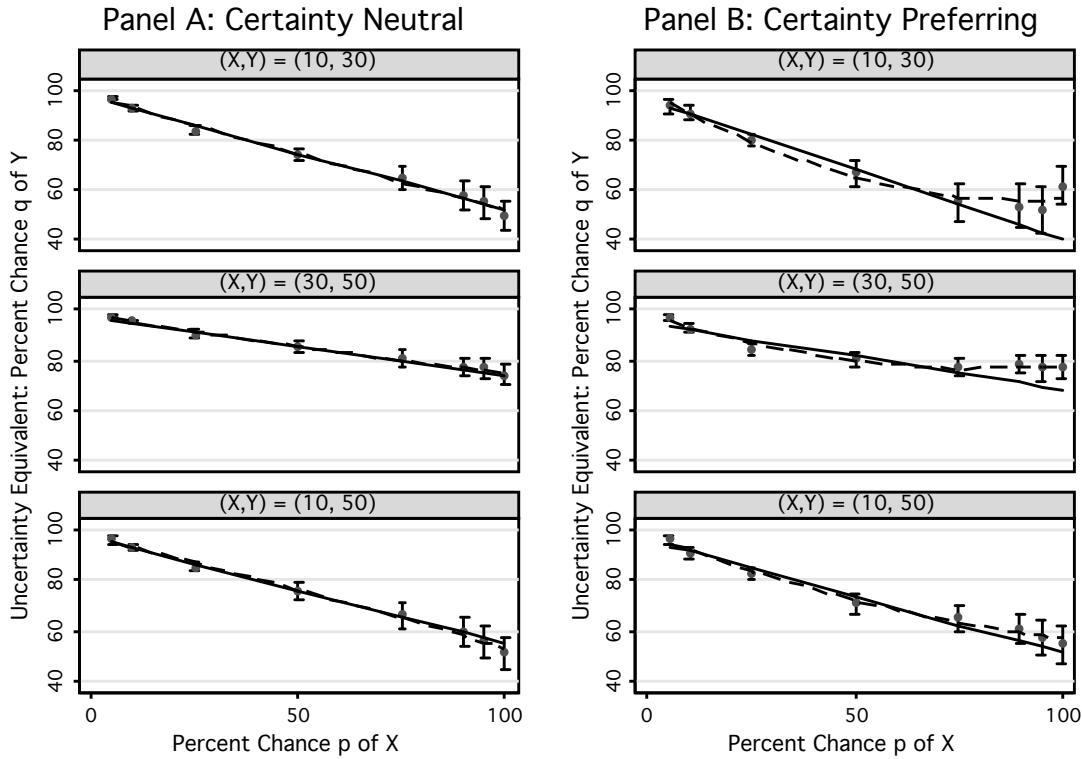
demonstrating that one cannot likely consider near-certainty dominance violations as an error and probability weighting as a true preference. The two phenomena correlate highly at the individual level.

³⁵This is not a complete taxonomy of types as one could imagine a classification for *Certainty Averse*. A full axiomatic development of Certainty Preferent, Neutral and Averse is left for future work and the present classifications are consistent with violation and non-violation of stochastic dominance between $p = 1$ and $p' = 0.95$. There were no session or order effects obtained for stochastic dominance violation rates or categorization of certainty preference. Certainty Preferring individuals are also more likely to violate stochastic dominance away from certainty. Their violation rate away from certainty is 8.2% (7.5%) versus 1.9% (4.1%) for Certainty Neutral subjects, ($t = 4.70$, $p < 0.001$). This, however, is largely driven by violations close to certainty.

³⁶See Appendix Tables A1 and A2 for full estimates.

natural $u-v$ specification one would expect less pronounced violations of stochastic dominance as experimental stakes diverge in value.³⁷

Figure 4: Uncertainty Equivalents and Certainty Preference



Note: Figure presents estimated uncertainty equivalent, $(q; Y, 0)$, for each given gamble, $(p; X, Y)$, of the experiment split by certainty preference, following methodology from Table 1, Panel A. Dashed black line represents the quadratic model fit following methodology from Table 1, Panel B. The solid black line corresponds to a linear projection based upon data from $p \leq 0.75$, indicating the degree to which the data adhere to the expected utility prediction of linearity away from certainty. See Appendix Tables A1 and A2 for estimates.

Our finding of *within-subject* violations of stochastic dominance is support for the hotly debated ‘uncertainty effect.’ Gneezy et al. (2006) discuss between-subject results indicating that a gamble over book-store gift certificates is valued less than the certainty of the gamble’s worst outcome. Though the effect was reproduced in Simonsohn

³⁷The Andreoni and Sprenger (2009c) specification is of $u-v$ preferences with $v(x) = x^\alpha$, $u(x) = x^{\alpha-\beta}$ with $\beta < \alpha < 1$. The differential curvature causes less pronounced violations of stochastic dominance when stakes differ substantially.

(2009), other work has challenged these results (Keren and Willemsen, 2008; Rydval et al., 2009). While Gneezy et al. (2006) do not find within-subject examples of the uncertainty effect, Sonsino (2008) finds a similar within-subject effect in the Internet auction bidding behavior of around 30% of individuals. Additionally, the uncertainty effect was thought not to be present for monetary payments (Gneezy et al., 2006). Our findings may help to inform the debate on the uncertainty effect and its robustness to the monetary domain. Additionally, our results may also help to identify the source of the uncertainty effect: a disproportionate preference for certainty. Indeed, this view is hypothesized by Gneezy et al. (2006), who suggest that “an individual posed with a lottery that involves equal chance at a \$50 and \$100 gift certificate might code this lottery as a \$75 gift certificate plus some risk. She might then assign a value to a \$75 gift certificate (say \$35), and then reduce this amount (to say \$15) to account for the uncertainty.” [p. 1291]

4.2 Certainty Equivalents Data

Seven certainty equivalents tasks with p gambles over \$30 and \$0 from the set $p \in \{0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95\}$ were administered, following the probabilities used in the original probability weighting experiments of Tversky and Kahneman (1992) and Tversky and Fox (1995). The analysis also follows closely the presentation and non-linear estimation techniques of Tversky and Kahneman (1992) and Tversky and Fox (1995).

As noted in Section 2, certainty equivalent analysis estimating risk aversion or probability weighting parameters that assumes a single utility function will be misspecified if there exists a disproportionate preference for certainty. As such, extreme small-stakes risk aversion or non-linear probability weighting may be apparent when none actually exists.³⁸ We first document small stakes risk aversion and apparent probability weight-

³⁸Diecidue et al. (2004) discuss potential functional forms that could deliver both apparent up-weighting of low probabilities and down-weighting of high probabilities in certainty equivalents. One

ing in our data and then correlate these phenomena with the violations of dominance measured in Section 4.1.

4.2.1 Risk Aversion and Probability Weighting

The identification of probability weighting and small-stakes risk aversion from certainty equivalents data normally relies on a range of experimental probabilities from near zero to near one. Probability weighting is initially supported if, for fixed stakes, subjects appear risk loving at low probabilities and risk averse at higher probabilities.³⁹ Small-stakes risk aversion would be viewed as the risk aversion aspect of this phenomenon.

Figure 5 presents a summary of the certainty equivalents.⁴⁰ As in sub-section 4.1, we first conducted an interval regression of the certainty equivalent, C , on indicators for the experimental probabilities (corresponding estimates are provided in Appendix Table A3, column 1). Following Tversky and Kahneman (1992), the data are presented relative to a benchmark of risk neutrality such that, for a linear utility function, Figure 5 directly reveals the probability weighting function, $\pi(p)$. The data show evidence of both small stakes risk aversion and non-linear probability weighting. Subjects appear significantly risk loving at low probabilities and significantly risk averse at intermediate and high probabilities. These findings are in stark contrast to those obtained in the uncertainty equivalents discussed in Section 4.1. Whereas in uncertainty equivalents we obtain no support for S -shaped probability weighting, in certainty equivalents we reproduce the probability weighting results generally found.⁴¹

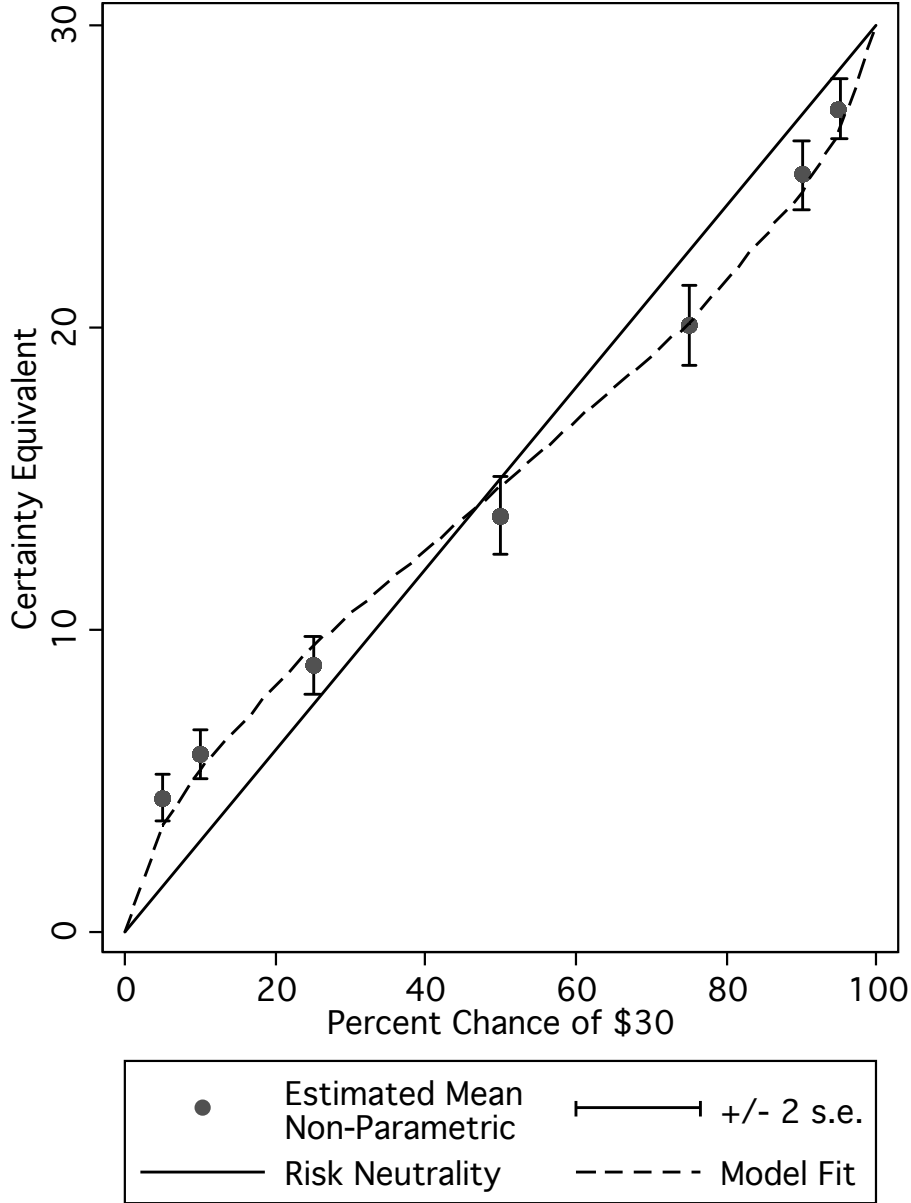
possibility is the u - v parameterization discussed in Andreoni and Sprenger (2009c) with differential curvature, $v(x) = x^\alpha$, $u(x) = x^{\alpha-\beta}$ with $\beta < \alpha < 1$, which produces both up-weighting of (very) low probabilities and down-weighting of high probabilities.

³⁹Because certainty equivalent responses are determined by both utility function curvature and probability weighting, even risk aversion at low probabilities could be consistent with probability weighting provided risk aversion was increasing in probability.

⁴⁰Figure 5 excludes one subject with multiple switching in one task. Identical aggregate results are obtained with the inclusion of this subject. However, we cannot estimate probability weighting at the individual level for this subject.

⁴¹Certainty equivalents correlate significantly with the number of safe choices in the Holt-Laury risk tasks. For example, for $p = 0.5$ the individual correlations between the midpoint certainty equivalent, C , and the number of safe choices, S_{10} and S_{30} , in the HL tasks are $\rho_{C,S_{10}} = -0.24$ ($p < 0.05$)

Figure 5: Certainty Equivalent Responses



Note: Mean certainty equivalent response. Solid line corresponds to risk neutrality. Dashed line corresponds to fitted values from non-linear least squares regression (1).

and $\rho_{C,S_{30}} = -0.24$ ($p < 0.05$). These results demonstrate consistency across elicitation techniques as a lower certainty equivalent and a higher number of safe HL choices both indicate more risk aversion. Additionally, the certainty equivalents correlate significantly with uncertainty equivalents. For example, for $p = 0.5$ the individual correlations between the midpoint certainty equivalent, C , and the midpoint of the uncertainty equivalent, q , are $\rho_{C,q(10,30)} = -0.24$ ($p < 0.05$), $\rho_{C,q(30,50)} = -0.25$ ($p < 0.05$), and $\rho_{C,q(10,50)} = -0.24$ ($p < 0.05$).

Tversky and Kahneman (1992) and Tversky and Fox (1995) obtain probability weighting parameters from certainty equivalents data by parameterizing both the utility and probability weighting functions and assuming the indifference condition

$$u(C) = \pi(p) \cdot u(30)$$

is met for each observation. We follow the parameterization of Tversky and Kahneman (1992) with power utility, $u(X) = X^\alpha$, and the one-parameter weighting function $\pi(p) = p^\gamma / (p^\gamma + (1 - p)^\gamma)^{1/\gamma}$.⁴² Lower γ corresponds to more intense probability weighting. The parameters $\hat{\gamma}$ and $\hat{\alpha}$ are then estimated as the values that minimize the sum of squared residuals of the non-linear regression equation

$$C = [p^\gamma / (p^\gamma + (1 - p)^\gamma)^{1/\gamma} \times 30^\alpha]^{1/\alpha} + \epsilon. \quad (1)$$

When conducting such analysis on our aggregate data with standard errors clustered on the subject level, we obtain $\hat{\alpha} = 1.07$ (0.05) and $\hat{\gamma} = 0.73$ (0.03).⁴³ The hypothesis of linear utility, $\alpha = 1$, is not rejected, ($F_{1,74} = 2.18$, $p = 0.15$), while linearity in probability, $\gamma = 1$, is rejected at all conventional levels, ($F_{1,74} = 106.36$, $p < 0.01$). The model fit is presented as the dashed line in Figure 5. The obtained probability weighting estimate compares favorably with the Tversky and Kahneman (1992) estimate of $\hat{\gamma} = 0.61$ and other one-parameter estimates such as Wu and Gonzalez (1996) who estimate $\hat{\gamma} = 0.71$.

⁴²Tversky and Fox (1995) use power utility with curvature fixed at $\alpha = 0.88$ from Tversky and Kahneman (1992) and a two parameter $\pi(\cdot)$ function.

⁴³For this analysis we estimate using the interval midpoint as the value of C , and note that the dependent variable is measured with error.

4.3 Reconciling Uncertainty and Certainty Equivalents

One hypothesis for the apparent presence of small stakes risk aversion and probability weighting in the aggregate certainty equivalents data is that they are the result of a bias due to misspecification. Could the estimates just presented be consistent with a disproportionate preference for certainty? In order to test this hypothesis, we define the variable *Violation Rate* as the stochastic dominance violation rate for the 21 comparisons involving certainty in the uncertainty equivalents.⁴⁴ Violation Rate is a continuous measure of the the degree to which individuals violate stochastic dominance at certainty and so a continuous measure of the intensity of the preference for certainty⁴⁵, which we correlate with certainty equivalents. In Figure 6 we find significant negative correlations between Violation Rate and certainty equivalents, primarily at higher probabilities. Insignificant positive correlations are found at lower probabilities. These results indicate that subjects with a more intense preference for certainty display significantly more small stakes risk aversion. These results are confirmed in regression, and we reject the null hypothesis that Violation Rate has no influence on certainty equivalent responses, ($\chi^2(7) = 18.06$, $p < 0.01$, Appendix Table A3, Column (5) provides the detail).

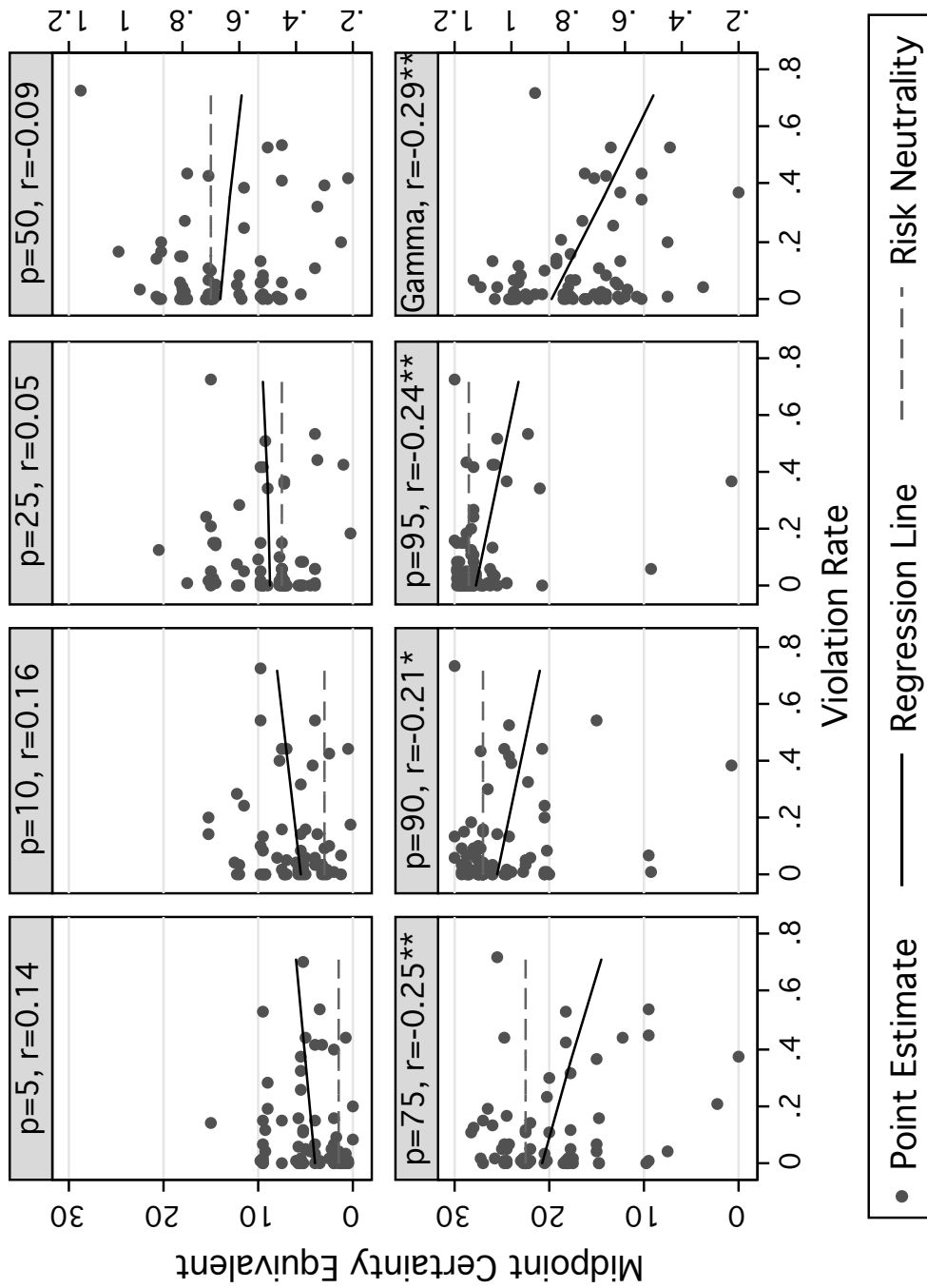
For subjects with a more intense preference for certainty, the significant increase in risk aversion at high probabilities and the slight increase in risk loving at low probabilities introduces more non-linearity into their estimated probability weighting functions. Figure 6 also presents the correlation between individual probability weighting, $\hat{\gamma}$, estimated from (1) and the intensity of certainty preference, Violation Rate.⁴⁶ The degree

⁴⁴A small minority of Certainty Neutral subjects have non-zero violation rates, as their elicited q at certainty is higher than that of some lower probability. The average certainty Violation Rate (0.069) for the 19% of Certainty Neutral subjects (9 of 47) with positive Violation Rate values is about same as their average violation rate away from certainty (0.060). For Certainty Preferring subjects, the average certainty Violation Rate (0.235) is about three times their violation rate away from certainty (0.082).

⁴⁵We recognize that this is a rough measure of intensity of certainty preference in the sense that individuals could have a non-monotonic relationship between p and q away from certainty. However, given the low dominance violation rates away from certainty, this is not overly problematic.

⁴⁶Following the aggregate estimate, $\alpha = 1$ is assumed for the individual estimates.

Figure 6: Certainty Equivalents (C), Probability Weighting ($\hat{\gamma}$), and Violation Rate



Note: Correlations (r) for 75 of 76 subjects. One subject with multiple switching in one certainty equivalent task not included. One percent jitter added to scatter plots. Two observations with Violation Rate = 0 and $\hat{\gamma} \geq 2$ not included in scatter plots, but included in regression line and reported correlation. Level of significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

to which individuals disproportionately prefer certainty predicts the degree of certainty equivalent-elicited probability weighting, ($\rho = -0.29$, $p = 0.011$).⁴⁷ This gives support to the claim that a disproportionate preference for certainty is conflated with non-linear probability weighting in standard certainty-based experiments.⁴⁸

It is critically important to contrast our results demonstrating the absence of probability weighting in uncertainty equivalents, its presence in certainty equivalents, and its relationship to violations of stochastic dominance with the body of results that find support for *S*-shaped probability weighting. Above we have discussed why standard estimation exercises assuming functional form for utility or probability weighting cannot be used to directly test linearity-in-probabilities. However, there exist a number of studies parametrically investigating probability weighting in decisions without certainty (Tanaka, Camerer and Nguyen, 2010; Booij, van Praag and van de Kuilen, 2010). These parametric estimates indicate that *S*-shaped probability weighting may be observed in decisions without certainty and clearly points to the need for future research. Additionally, attention must be given to the ‘parameter-free’ elicitation techniques that find non-parametric support for non-linear probability weights (Gonzalez and Wu, 1999; Abdellaoui, 2000; Bleichrodt and Pinto, 2000). Importantly, both Gonzalez and Wu (1999) and Abdellaoui (2000) make use of certainty equivalents or a number of certain outcomes to identify probability weights, a technique that is mis-

⁴⁷Additionally, the indicator for Certainty Preferent correlates with certainty equivalent response. See Appendix Table A3, Column (3). Eleven subjects exhibited non-monotonic relationships between experimental probabilities and elicited certainty equivalents. Though no systematic pattern was observed, such behavior is another example of violating stochastic dominance in decisions involving certainty. Non-monotonicity correlates significantly with being Certainty Preferent, ($\rho = 0.44$, $p < 0.01$). Eliminating these individuals leaves the results qualitatively unchanged. Certainty equivalent behavior remains supportive of probability weighting and elicited probability weights remain significantly correlated with Violation Rate, ($\rho = -0.29$, $p < 0.05$). Appendix Table A3, Columns (2), (4) and (6) provide estimates.

⁴⁸As a robustness test, we repeat the analysis with an alternate functional form Prelec (1998), $\pi(p) = \exp(-(-\ln p)^\gamma)$, and obtain a correlation between $\hat{\gamma}$ and Violation Rate of $\rho = -0.25$ $p < 0.05$. Additionally, Appendix Tables A4 and A5 present results from the classification of risk attitudes based on the interval of the certainty equivalent response. These data demonstrate that Certainty Preferring subjects are more likely to be classified as risk loving at low probabilities and are more likely to be classified as risk averse at higher probabilities. Additionally, Certainty Neutral subjects are more likely to be risk neutral at higher probabilities.

specified if there exists a specific preference for certainty. Bleichrodt and Pinto (2000) do not use certainty equivalents techniques, but their experiment is designed not to elicit preferences over monetary payments, but rather over hypothetical life years. It is not clear the extent to which such findings apply to incentivized elicitation procedures over money.⁴⁹

5 Conclusion

Volumes of research exists exploring both the implications and violations of the independence axiom. Surprisingly, little research exists directly testing the most critical result of the independence axiom: linearity-in-probabilities of the Expected Utility (EU) function. We present an experimental device that easily generates such a direct test, the *uncertainty equivalent*. Uncertainty equivalents not only provide tests of expected utility's linearity-in-probabilities, but also provide separation between competing alternative preference models such as Cumulative Prospect Theory's (CPT) inverted *S*-shaped probability weighting (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992), expectations-based reference dependence (Bell, 1985; Loomes and Sugden, 1986; Gul, 1991; Koszegi and Rabin, 2006, 2007), and the *u-v* preference model (Neilson, 1992; Schmidt, 1998; Diecidue et al., 2004).

In a within-subject experiment with both uncertainty equivalents and standard certainty equivalent methodology we demonstrate four important results. First, independence performs well away from certainty where probabilities are found to be weighted nearly linearly. Second, independence breaks down close to certainty. The nature of the violation is contrary to standard *S*-shaped probability weighting and consistent with other alternative models such as disappointment aversion and *u-v* preferences,

⁴⁹Abdellaoui (2000), Bleichrodt and Pinto (2000), Booij and van de Kuilen (2009) and Booij et al. (2010) share a two-stage elicitation procedure which 'chains' responses in order to obtain utility or probability weighting values. Such chained procedures are common to the 'trade-off' (Wakker and Deneffe, 1996) method of utility assessment. A discussed problem with these chained methods is that errors propagate through the experiment.

which both feature a disproportionate preference for certainty. Third, nearly 40% of experimental subjects indirectly violate first order stochastic dominance as probabilities approach 1. These violations are a necessary prediction of the $u-v$ model and are accommodated in some versions of disappointment aversion. Fourth, in certainty equivalents experiments, apparent S -shaped probability weighting and small stakes risk aversion phenomena are observed, closely reproducing prior findings. However, these phenomena are driven by individuals who exhibit a disproportionate preference for certainty in uncertainty equivalents by violating stochastic dominance.

Our findings have critical implications for research on risk attitudes and have applications to a variety of economic problems. The results demonstrate that experimental measures of risk attitudes and EU violations are dramatically influenced by the presence of certainty. Since the work of Allais (1953b) certainty has been known to play a special role in decision-making and in generating non-EU behavior. Our results indicate that a specific preference for certainty may be the key element in producing such behavior. This suggests that empirical work should take great care to separate certainty preferences from other phenomena under investigation. Additionally, theoretical research should take seriously models with specific preferences for certainty, (small scale) violations of stochastic dominance, and their implications for decision-making under uncertainty.

References

- Abdellaoui, Mohammed**, “Parameter-Free Elicitation of Utility and Probability Weighting Functions,” *Management Science*, 2000, *46* (11), 1497–1512.
- Abeler, Johannes, Armin Falk, Lorenz Goette, and David Huffman**, “Reference Points and Effort Provision,” *American Economic Review*, Forthcoming.
- Allais, Maurice**, “Fondements d’une Théorie Positive des Choix Comportant un Risque,” in “Colloques Internationaux du Centre National de la Recherche Scientifique (Econometrie) 40” 1953, pp. 127–140.
- , “Le Comportement de l’Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l’Ecole Americaine,” *Econometrica*, 1953, *21* (4), 503–546.
- Andreoni, James and Charles Sprenger**, “Certain and Uncertain Utility: The Allais Paradox and Five Decision Theory Phenomena,” *Working Paper*, 2009c.
- and —, “Risk Preferences Are Not Time Preferences,” *Working Paper*, 2011.
- Bell, David E.**, “Disappointment in Decision Making under Uncertainty,” *Operations Research*, 1985, *33* (1), 1–27.
- Bernouilli, Daniel**, “Specimen Theoriae Novae de Mensura Sortis,” *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, 1738, *5*, 175–192.
- Birnbaum, Michael H.**, “Violations of Monotonicity and Contextual Effects in Choice-Based Certainty Equivalents,” *Psychological Science*, 1992, *3*, 310–314.
- Bleichrodt, Han and Jose Luis Pinto**, “A Parameter-Free Elicitation of the Probability Weighting Function in Medical Decision Analysis,” *Management Science*, 2000, *46* (11), 1485–1496.
- , **Jose Maria Abellan-Perinan, Jose Luis Pinto-Prades, and Ildefonso Mendez-Martinez**, “Resolving Inconsistencies in Utility Measurement Under Risk: Tests of Generalizations of Expected Utility,” *Management Science*, 2007, *53* (3).
- Booij, Adam S. and Gijs van de Kuilen**, “A Parameter-Free Analysis of the Utility of Money for the General Population Under Prospect Theory,” *Journal of Economic Psychology*, 2009, *30*, 651–666.
- , **Bernard M. S. van Praag, and Gijs van de Kuilen**, “A Parametric Analysis of Prospect Theory’s Functionals for the General Population,” *Theory and Decision*, 2010, *68*, 115–148.
- Camerer, Colin F.**, “Recent Tests of Generalizations of Expected Utility Theory,” in Ward Edwards, ed., *Utility: Theories, Measurement, and Applications*, Kluwer: Norwell, MA, 1992, pp. 207–251.
- Conlisk, John**, “Three Variants on the Allais Example,” *The American Economic Review*, 1989, *79* (3), 392–407.

- Diecidue, Enrico, Ulrich Schmidt, and Peter P. Wakker**, “The Utility of Gambling Reconsidered,” *Journal of Risk and Uncertainty*, 2004, 29 (3), 241–259.
- Epstein, Larry G. and Stanley E. Zin**, “Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework,” *Econometrica*, 1989, 57 (4), 937–969.
- Ericson, Keith M. Marzilli and Andreas Fuster**, “Expectations as Endowments: Reference-Dependent Preferences and Exchange Behavior,” *The Quarterly Journal of Economics*, Forthcoming.
- Farquhar, Peter H.**, “Utility Assessment Methods,” *Management Science*, 1984, 30 (11), 1283–1300.
- Fishburn, Peter and Peter Wakker**, “The Invention of the Independence Condition for Preferences,” *Management Science*, 1995, 41 (7), 1130–1144.
- Gill, David and Victoria Prowse**, “A Structural Analysis of Disappointment Aversion in a Real Effort Competition,” *Working Paper*, 2010.
- Gneezy, Uri, John A. List, and George Wu**, “The Uncertainty Effect: When a Risky Prospect Is Valued Less Than Its Worst Possible Outcome,” *The Quarterly Journal of Economics*, 2006, 121 (4), 1283–1309.
- Gonzalez, Richard and George Wu**, “On the Shape of the Probability Weighting Function,” *Cognitive Psychology*, 1999, 38, 129–166.
- Gul, Faruk**, “A Theory of Disappointment Aversion,” *Econometrica*, 1991, 59 (3), 667–686.
- Harless, David W. and Colin F. Camerer**, “The Predictive Utility of Generalized Expected Utility Theories,” *Econometrica*, 1994, 62 (6), 1251–1289.
- Harrison, Glenn W. and Elisabet E. Rutstrom**, “Risk Aversion in the Laboratory,” in Glenn W. Harrison and James C. Cox, eds., *Research in Experimental Economics: Volume 12. Risk Aversion in Experiments*, Bingley: Emerald, 2008.
- , **Morten I. Lau, Elisabet E. Rutstrom, and Melonie B. Williams**, “Eliciting risk and time preferences using field experiments: Some methodological issues,” in Jeffrey Carpenter, Glenn W. Harrison, and John A. List, eds., *Field experiments in economics*, Vol. Vol. 10 (Research in Experimental Economics), Greenwich and London: JAI Press, 2005.
- Hey, John D. and Chris Orme**, “Investigating Generalizations of Expected Utility Theory Using Experimental Data,” *Econometrica*, 1994, 62 (6), 1291–1326.
- Holt, Charles A. and Susan K. Laury**, “Risk Aversion and Incentive Effects,” *The American Economic Review*, 2002, 92 (5), 1644–1655.
- Jacobson, Sarah and Ragan Petrie**, “Learning from Mistakes: What Do Inconsistent Choices Over Risk Tell Us?,” *Journal of Risk and Uncertainty*, 2009, 38 (2).

- Kachelmeier, Steven J. and Mahamed Shehata**, “Examining Risk Preferences under Higher Monetary Incentives: Experimental Evidence from the People’s Republic of China,” *American Economic Review*, 1992, *82* (2), 1120–1141.
- Kahneman, Daniel and Amos Tversky**, “Prospect Theory: An Analysis of Decision under Risk,” *Econometrica*, 1979, *47* (2), 263–291.
- Keren, Gideon and Martijn C. Willemsen**, “Decision Anomalies, Experimenter Assumptions and Participants’ Comprehension: Revaluating the Uncertainty Effect,” *Journal of Behavioral Decision Making*, 2008, *22* (3), 301–317.
- Koszegi, Botond and Matthew Rabin**, “A Model of Reference-Dependent Preferences,” *Quarterly Journal of Economics*, 2006, *121* (4), 1133–1165.
- and —, “Reference-Dependent Risk Attitudes,” *The American Economic Review*, 2007, *97* (4), 1047–1073.
- Kreps, David M. and Evan L. Porteus**, “Temporal Resolution of Uncertainty and Dynamic Choice Theory,” *Econometrica*, 1978, *46* (1), 185–200.
- Loomes, Graham and Robert Sugden**, “Disappointment and Dynamic Consistency in Choice under Uncertainty,” *Review of Economic Studies*, 1986, *53* (2), 271–82.
- Magat, Wesley A., W. Kip Viscusi, and Joel Huber**, “A Reference Lottery Metric for Valuing Health,” *Management Science*, 1996, *42* (8), 1118–1130.
- Malinvaud, Edmond**, “Note on von Neumann-Morgenstern’s Strong Independence,” *Econometrica*, 1952, *20* (4), 679.
- McCord, Mark and Richard de Neufville**, “‘Lottery Equivalents’: Reduction of the Certainty Effect Problem in Utility Assignment,” *Management Science*, 1986, *32* (1), 56–60.
- Meier, Stephan and Charles Sprenger**, “Present-Biased Preferences and Credit Card Borrowing,” *American Economic Journal - Applied Economics*, 2010, *2* (1), 193–210.
- Neilson, William S.**, “Some Mixed Results on Boundary Effects,” *Economics Letters*, 1992, *39*, 275–278.
- Oliver, Adam**, “Testing the Internal Consistency of the Lottery Equivalents Method Using Health Outcomes,” *Health Economics*, 2005, *14*, 149–159.
- , “A Qualitative Analysis of the Lottery Equivalents Method,” *Economics and Philosophy*, 2007, *23*, 185–204.
- Prelec, Drazen**, “The Probability Weighting Function,” *Econometrica*, 1998, *66* (3), 497–527.
- Quiggin, John**, “A Theory of Anticipated Utility,” *Journal of Economic Behavior and Organization*, 1982, *3*, 323–343.
- Rabin, Matthew**, “Risk aversion and expected utility theory: A calibration theorem,” *Econometrica*, 2000a, *68* (5), 1281–1292.

- , “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, pp. 202–208.
- Rydval, Ondrej, Andreas Ortmann, Sasha Prokosheva, and Ralph Hertwig**, “How Certain is the Uncertainty Effect,” *Experimental Economics*, 2009, *12*, 473–487.
- Samuelson, Paul A.**, “Probability, Utility, and the Independence Axiom,” *Econometrica*, 1952, *20* (4), 670–678.
- , “Utilité, Préférence et Probabilité,” in “Colloques Internationaux du Centre National de la Recherche Scientifique (Econometrie) 40” 1953, pp. 141–150.
- Savage, Leonard J.**, “Une Axiomatisation de Comportement Raisonnable Face à l’Incertitude,” in “Colloques Internationaux du Centre National de la Recherche Scientifique (Econometrie) 40” 1953, pp. 29–33.
- , *The Foundations of Statistics*, New York: J. Wiley, 1954.
- Schmidt, Ulrich**, “A Measurement of the Certainty Effect,” *Journal of Mathematical Psychology*, 1998, *42* (1), 32–47.
- Schoemaker, Paul J. H.**, “The Expected Utility Model: Its Variants, Purposes, Evidence and Limitations,” *Journal of Economic Literature*, 1982, *20* (2), 529–563.
- Simonsohn, Uri**, “Direct Risk Aversion: Evidence from Risky Prospects Valued Below Their Worst Outcome,” *Psychological Science*, 2009, *20* (6), 686–692.
- Sonsino, Doron**, “Disappointment Aversion in Internet Bidding Decisions,” *Theory and Decision*, 2008, *64* (2-3), 363–393.
- Sprenger, Charles**, “An Endowment Effect for Risk: Experimental Tests of Stochastic Reference Points,” *Working Paper*, 2010.
- Starmer, Chris**, “Developments in Non-Expected Utility Theory: The Hunt for a Descriptive Theory of Choice Under Risk,” *Journal of Economic Literature*, 2000, *38* (2).
- **and Robert Sugden**, “Does the Random-Lottery Incentive System Elicit True Preferences? An Experimental Investigation,” *The American Economic Review*, 1991, *81* (4), 971–978.
- Stewart, Mark B.**, “On Least Squares Estimation when the Dependent Variable is Grouped,” *The Review of Economic Studies*, 1983, *50* (4), 737–753.
- Tanaka, Tomomi, Colin Camerer, and Quang Nguyen**, “Risk and time preferences: Experimental and household data from Vietnam,” *American Economic Review*, 2010, *100* (1), 557–571.
- Tversky, Amos and Craig R. Fox**, “Weighing Risk and Uncertainty,” *Psychological Review*, 1995, *102* (2), 269–283.
- **and Daniel Kahneman**, “Advances in Prospect Theory: Cumulative Representation of Uncertainty,” *Journal of Risk and Uncertainty*, 1992, *5* (4), 297–323.

Varian, Hal R., *Microeconomic Analysis*, 3rd ed., New York: Norton, 1992.

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

Wakker, Peter P. and Daniel Deneffe, “Eliciting von Neumann-Morgenstern Utilities When Probabilities are Distorted or Unknown,” *Management Science*, 1996, 42 (8), 1131–1150.

Wu, George and Richard Gonzalez, “Curvature of the Probability Weighting Function,” *Management Science*, 1996, 42 (12), 1676–1690.