7-1-2016

Cumulative Prospect Theory in the Laboratory: A Reconsideration

Glenn Harrison
Georgia State University

Todd Swarthout
University of Cape Town

Cumulative Prospect Theory in the Laboratory: A Reconsideration

by

Glenn W. Harrison and J. Todd Swarthout †

July 2016

ABSTRACT.

We take Cumulative Prospect Theory (CPT) seriously by rigorously estimating structural models using the full set of CPT parameters. Much of the literature only estimates a subset of CPT parameters, or more simply assume CPT parameter values from prior studies. Our data are from substantial laboratory experiments with undergraduate students and MBA students facing real incentives and losses. We also estimate structural models from Expected Utility Theory, Dual Theory, Rank-Dependent Utility (RDU) and Disappointment Aversion for comparison. Our major finding is that a majority of individuals in our sample locally asset integrate. That is, they see a loss frame for what it is, a frame, and behave as if they evaluate the net payment rather than the gross loss when one is presented to them. This finding is devastating to the *direct* application of CPT to these data for those subjects. Support for CPT is greater when losses are covered out of an earned endowment rather than house money, but RDU is still the best single characterization of individual and pooled choices. Defenders of the CPT model claim, correctly, that the CPT model exists "because the data says it should." In other words, the CPT model was borne from a wide range of stylized facts culled from parts of the cognitive psychology literature. If one is to take the CPT model seriously and rigorously then it needs to do a much better job of explaining the data than we see here.

† Department of Risk Management & Insurance and Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University, USA (Harrison); and Department of Economics, Andrew Young School of Policy Studies, Georgia State University, USA (Swarthout). Harrison is also affiliated with the School of Economics, University of Cape Town. E-mail contacts: gharrison@gsu.edu and swarthout@gsu.edu. We are grateful to *Bloomberg Wealth* for funding.

Table of Contents

Is the empirical evidence for Cumulative Prospect Theory as strong as it is claimed to be? We go back to the laboratory[1] to undertake simple, direct tests of the hypothesis that individuals make choices over risky lotteries in the manner assumed by Cumulative Prospect Theory (CPT). We go back to the "base camp" of evidence for CPT from Tversky and Kahneman [1992] and evaluate parametric, structural models for individuals making incentivized choices over risky lotteries defined over the gain frame, the loss frame, and the mixed frame. Our results are surprising.

We focus on five core models of decision-making under objective risk. One is Expected Utility Theory (EUT), and posits that the risk premium is explained solely by an aversion to variability of earnings from a prospect. The second model is Dual Theory (DT) which assumes no aversion to variability of earnings from a prospect, but instead posits that decision-makers may be pessimistic or optimistic with respect to the probabilities of outcomes. The third model is Rank-Dependent Utility (RDU), which allows both of the latent psychological processes of EUT and DT in combination. RDU does not rule out aversion to variability of earnings, as in DT, but just augments it with an additional latent psychological process. The fourth model is Disappointment Aversion (DA), which assumes that individuals evaluate prospects according to an augmented version of EUT, in which they also take into account the extent to which outcomes differ from the certainty-equivalent of the prospect. EUT, DT, RDU and DA assume that individuals asset integrate, in the sense that they net out framed losses from some endowment. The final model is CPT, which adds to RDU an aversion to losses as a possible psychological pathway to the risk premium, and also adds the assumption that gross gains and losses matter because individuals do not

[1] We appreciate that there is a substantial literature claiming that CPT, and specifically loss aversion, is well-documented in the field. We are staunch advocates of the value of field experiments, as in Harrison and List [2004], but only as a complement to what we can more efficiently learn from the laboratory. For instance, Ray, Shum and Camerer [2015; p.376] note that loss aversion "... was originally discovered in laboratory choices among monetary risks [...], but has since become evident in allocations, behavior, and institutional rules in many markets." Given the paucity of evidence for CPT in the controlled setting of the laboratory when we demand the use of real incentives, experimental designs that allow the conceptual identification of loss aversion, and explicit structural estimation, we are skeptical about claims from the field, where there are always acknowledged confounds to clean inference.

locally asset integrate and evaluate net gains or losses.

We find that the vast majority of individuals in our sample appear to locally asset integrate. That is, they see the loss frame for what it is, a frame, and behave as if they evaluate the net payment rather than the gross loss when one is presented to them. This finding is devastating to the direct application of CPT to these data. We find greater support for RDU than CPT, and in many cases greater support for EUT over CPT. We find virtually no support for DT or DA. At the individual level, almost all of our subjects can be classified using EUT and RDU, with a majority being RDU.

In section 1 we outline the experimental design we developed, in which each subject is given 100 binary lottery choices defined over the gain frame, loss frame and mixed frame. The parameter values were designed, following a neglected design of Loomes and Sugden [1998], to provide stress tests of the independence axiom of EUT as well as to allow for identification of a wide range of risk attitudes. In section 2 we lay out the models to be estimated, with particular care over the specification of the CPT model. A by-product is attention to the theoretical implications for the CPT model, implications for experimental design, and a detailed statement of the manner in which mixed frame lotteries are handled. All of these issues are "in the literature," but scattered and often neglected. In section 3 we present results. Section 4 seeks to connect our approach and findings to the vast literature, noting the remarkably under-developed state of structural estimation for the CPT model once we apply minimal methodological requirements. We discuss limitations of our results in Section 5, and offer conclusions in Section 6, focusing on the variants that should be examined next in the rigorous evaluation of CPT.

1. Experiments

Our objective is to design a battery of tests that allows identification of all of the parameters of the EUT, DT, RDU, DA and CPT models, that provides some "stress tests" of the EUT model, and that allows estimation of a wide range of risk preferences at the individual level. The first criterion means that we must have gain frame lotteries, loss frame lotteries, and mixed frame lotteries. The terminology "gain" and "loss" refer here to lotteries in which all prizes are (weakly) gains or losses, and the terminology "mixed" refers here to lotteries in which some prizes are (strictly) gains *and* some are (strictly) losses. The second criterion means that we need to present some *sets* of choices that generate sharp predictions under EUT, such as the classic Allais Paradox set of two choices, and the classic Common Ratio set of two choices. The third criterion means that we need to recognize that certain risk preferences could make individuals indifferent between the two lotteries in any given choice, and hence generate low power tests of EUT, DT or RDU. And it also means that we should try to generate stakes that are as large as possible, within obvious feasibility constraints for budgets. Following the vast literature, we focus on binary lottery choices, with a standard interface illustrated in Figure 1.

Some of the most important batteries of tests do not satisfy all three of these, nor were they designed to do so. For instance, the justifiably influential battery developed by Hey and Orme [1994] does not have loss or mixed frames, and deliberately avoided sets of lottery pairs that had generated "knife-edge" tests of EUT. Their design mantra was to be agnostic about choice patterns, and see which models best characterized the data, rather than selecting lottery pairs designed to be hard for EUT *per se*.

Loomes and Sugden [1998] pose an important design feature for Common Ratio tests, allowing us to meet the last two criteria: variation in the "gradient" of the EUT-consistent indifference curves within a Marschak-Machina (MM) triangle. The reason for this is variation is to generate some choice patterns that are more powerful tests of EUT for any given risk attitude.

Under EUT the slope of the indifference curve within a MM triangle is a measure of risk aversion. So there always exists some risk attitude such that the subject is indifferent, as stressed by Harrison [1984], and evidence of Common Ratio violations has virtually zero power.[2] This design can be visualized instantly from Loomes and Sugden [1998; Figure 2, p.587]. For our batteries of 100 lottery choices, explained in more detail below, the comparable MM visualizations are shown in Figures 2 and 3.

Each panel within Figures 2 and 3 refers to a specific context of prizes. The top left panel of Figure 2 has prizes of \$0, \$5 and \$25, and the panel immediately to its right has prizes of \$0, \$15 and \$75. There are always one, two or three prizes in each lottery that have positive probability of occurring. The vertical axis in each panel shows the probability attached to the high prize of that triple, and the horizontal axis shows the probability attached to the low prize of that triple. When the probability of the highest and lowest prize is zero, then 100% weight falls on the middle prize. This specific lottery is illustrated in the bottom left corner of the very first panel, where the subject has a lottery offering \$5 for certain. Any lotteries strictly in the interior of the MM triangle have positive weight on all three prizes, and any lottery on the boundary of the MM triangle has zero weight on one or two prizes.

The solid dots within each panel of Figures 2 and 3 show specific lotteries offered to subjects, and the lines show choice pairs offered. The detailed numerical patterns are listed in Appendix A. For the top left panel of Figure 2, we have the familiar Allais Paradox defined over real monetary outcomes. The lottery pair given by the chord in the bottom left corner has subject choose between {\$0, 0; \$5, 1; \$25, 0} and {\$0, 0.01; \$5, 0.89; \$25, 0.1}, and the lottery pair given by the chord in the bottom right corner has the subject choose between {\$0, 0.89; \$5, 0.11; \$25, 0} and

[2] EUT does not, then, predict 50:50 choices, as some casually claim. It does say that the expected utility differences will not explain behavior, and that then allows all sorts of psychological factors to explain behavior. In effect, EUT has *no* prediction in this instance, and that is not the same as predicting an even split.

{$0, 0.9; $5, 0.11; $25, 0.1}. Since the slopes of these two chords are the same, and we know that indifference curves under EUT are straight lines within the MM triangle, we can easily see the choice pattern predicted by EUT: pick the first lottery in each pair or pick the second lottery in each pair. EUT is violated if the first (second) lottery is picked in the first pair, and the second (first) lottery is picked in the second pair.

The difference between the first two panels of the top row of Figures 2 and 3 is simply the change in the prize context. The first panel shows a "low stakes" context, and the second panel shows a "high stakes" context.

Figure 2 shows the 100 lottery pairs presented to a sample of 177 undergraduate students sampled from the Georgia State University population, and Figure 3 shows the 100 lottery pairs presented to a sample of 94 MBA students sampled from the Georgia State University population. The only difference between Figures 2 and 3 are the prizes, with the domain of *net* prizes for the undergraduates spanning $0 up to $70, and spanning $0 up to $750 for the MBA students. We deliberately had a number of prize contexts for the MBA students that were identical to the domain of prizes that the undergraduates faced, so we could ascertain the pure effect of stake size. In fact, the common lotteries themselves were identical: these 24 common choices are in the 2 panels in the bottom row of Figures 2 and 3, and in the third panel of the penultimate row of Figures 2 and 3.

Comparing panels within Figures 2 or 3, the logic of the Loomes and Sugden [1998] design appears: to have several choices within a MM triangle that allow tests of EUT, but to vary the slope of the chord connecting lottery pairs. For instance, consider the four panels within Figure 3 with prizes of $0, $35 and $70 (the last panel of the first row, the second row, and the first panel of the third row). Each panel contains several tests of the Common Ratio effect conditional on a given risk attitude under EUT. Then the gradients change from panel to panel, implying that one should have one panel, and probably several, that provide more powerful tests of EUT for any given risk attitude than other panels. A subject might be indifferent between the choices in one panel of these four, but then that subject must, *by design*, have strict preferences for some or all of the other panels. Harrison, Johnson, McInnes and Rutström [2007] refer to this as a "complementary slack experimental design," since low-power tests of EUT in one panel mean that there *must* be higher-power tests of EUT in another panel.[3]

In our battery, 96 of the choice pairs were derived from the Loomes and Sugden [1998] logic. Our modest contribution is to have some prize contexts in which endowments were given to subjects and the prizes framed as losses. The result is that we have 16 lottery pairs with loss frames, and 16 lottery pairs with mixed frames.

We added two Allais Paradox pairs to this set of lottery choices. Conlisk [1989] presents a real version of the Allais Paradox, with the binary choices marked "Allais low" in Figures 2 and 3 (the very first panel). When subjects are presented with just one lottery choice, and he compares patterns of choices on a between-subjects basis, he finds no evidence whatsoever of the Allais Paradox. For some reason, this finding does not stop many from referring to the Allais Paradox as a well-known pattern of EUT violation.[4] Starmer [1992] provided one of the first explorations of the generality of the common consequence effect, concluding (p. 829) that "... if we wish to use experimental evidence as a basis for developing new theories of choice under uncertainty, we may

---

[3] The experimental setting in which they proposed this terminology is when one conducts Common Ratio tests for the usual gradients and also conducts Preference Reversal tests for the usual lottery patterns. The former tests generate indifference or low power for modest and widely observed levels of risk aversion, and the latter tests generate indifference or low power for risk neutral subjects. If one can identify, to some statistical tolerance, whether a subject is risk neutral or not, conducting both types of tests allows one to know *a priori* that *one* of the tests should be a more powerful test of EUT. Of course, no single test can be a powerful test of EUT for all subjects, if one allows a mix of risk averse and risk neutral subjects. The basic idea was clearly stated in Loomes and Sugden [1998; p. 589, especially fn. 3].

[4] This result is acknowledged by some, but not welcome. For instance, Cubitt, Starmer and Sugden [1989; p. 130] comment that these results are "... sometimes quoted as evidence that violations of EUT are less frequent in single choice than in random lottery designs. Conlisk investigated the Common Consequence effect using a single choice design. In each of the two relevant tasks, almost all subjects (26 out of 27 in one case, 24 out of 26 in the other) chose the riskier option. Clearly, this distribution of responses between riskier and safer choices is far too asymmetric for the experiment to be a satisfactory test for systematic deviations from EUT." The logic of the final sentence is hard to ascertain. Moreover, the evidence for the Common Consequence effect in *incentivized* "random lottery designs" is decidedly mixed: Burke, Carter, Gominiak and Ohl [1996] and Fan [2002] find no evidence of an EUT violation, whereas Starmer and Sugden [1991] do.

have to accept that the behavior of individuals is rather more subtle and complex than we have previously thought." We therefore include two instances of Allais Paradox tasks in our battery.

Figure 4 displays the complete set of probability patterns for our battery, ignoring the prize context. The rich array of slopes for the choice chords allows one to see why this design should provide an attractive setting to estimate models of risky choice, certainly from the perspective of stress-testing EUT against RDU and CPT.[5] For the undergraduate sample, average payouts for a risk-neutral subject would be just over $39, comparable to earlier experiments with this population, but on the high side historically for tests of EUT with real consequences. We consider the effect of prize "context" by using different prizes and the lottery choices in 24 gain frame choices. This allows us to move from having choices defined over prizes of only $0, $35 and $70, to having prizes of $0, $5, $10, $15, $20, $25, $30, $35, $45, $55, $60 and $70. This design feature will help estimation of a more precise utility function.

The lotteries for the MBA students are qualitatively identical to those for the undergraduate sample, but many are scaled up in monetary value by a multiplicative factor of 25. In addition, subjects were offered a healthy $40 show-up fee, to reflect the increased opportunity cost of their time compared to our convenience sample of GSU undergraduates. We scaled up 71 of the tasks, including all loss and mixed frame tasks, by 25 such that the prize *frame* varies from +$500 to -$500. One of the Allais Paradox tasks was scaled up by a factor of 20, so that the higher-stake prizes for this task are $0, $100 and $500. The low-stakes Allais Paradox task, and 24 gain-frame tasks, were left at the original scaling to allow comparability of behavior across the samples, and to assess the

[5] We use all 40 of the Loomes and Sugden [1998] choice patterns, given by triangles I through V of their design. Triangle VI was used for their £20 sample, and triangle V for their £30 sample. These provide us with 40 gain-frame choice pairs. We then use the choices from their triangles II and V for 16 loss-frame lottery pairs, and 16 mixed-frame lottery pairs. Triangle II shows considerable evidence of their subjects picking the "safe" lottery, so this will test if loss frames induce risk-seeking: there is somewhere for the choices to go, compared to the gain frame expectations. Triangle V has a mix of safe and risky choices, just slightly dominated by risky choices, so it provides a non-extreme baseline to see the effect of loss frames. Triangle V is also likely to be more statistically information than taking triangles III or IV on top of triangle II, due to the change in gradient.

effect of scaling up payoffs substantially. A risk-neutral individual would expect to earn $210.37 from these choices, before the show-up fee.

Appendix C contains the instructions given to subjects. Subjects had no other salient task in the experiment, although they did have to answer a series of hypothetical questions about "risk tolerance" after making their choices. We also collected standard demographic information.

We opted for using the Random Lottery Incentive Method (RLIM), where one of the 100 choices was to be chosen at random for playing out and payment. We did so for two reasons, recognizing that this is not as innocent a procedure as some maintain.[6] The first reason was to ensure that we collected choices over a wide enough array of lotteries to be able to identify the three competing models. If we had opted for giving one choice to each subject, to avoid using the RLIM, this would have been infeasible. The second reason was to be able to estimate at the level of the individual, to compare those estimates to pooled estimates over all individuals. Again, this would have been infeasible if we had given each subject just one choose.[7]

Total payouts from these experiments amounted to $42,258.

[6] See Cox, Sadiraj and Schmidt [2015] and Harrison and Swarthout [2014] for detailed arguments and evidence, as well as extensive literature discussion.

[7] The use of the RLIM is problematic, for well known theoretical reasons (Holt [1986] and Karni and Safra [1987]). Its use entails a certain "bipolar hypothesis" about the independence axiom when estimating RDU or CPT models: that the axiom works as it should when subjects evaluate the compound lottery over 100 simple lotteries that is implied by payment protocol, but that it magically blows up when subjects evaluate each simple lottery in the choice pair. One has to be simultaneously depressed and optimistically manic about the independence axiom to maintain these two positions. Moreover, Harrison and Swarthout [2015] find evidence that it *does* make a difference behaviorally when estimating RDU models, but not, as one would expect, when estimating EUT models. One logical response to this problem is just to assume two independence axioms: one axiom that applies to the evaluation of a given prospect, and that is assumed to be violated by DT, DA, RDU and CPT, and another axiom that applies to the evaluation of the experimental payment protocol. One can then assume failure of the former axiom, when estimating non-EUT models, but validity of the latter axiom.

2. Theoretical Models

A. Expected Utility Theory

Assume that utility of income is defined by

$$U(x) = x^{(1-r)}/(1-r) \tag{1}$$

where x is the lottery prize and $r \neq 1$ is a parameter to be estimated. For r=1 assume U(x)=ln(x) if needed. Thus r is the coefficient of CRRA: r=0 corresponds to risk neutrality, r<0 to risk loving, and r>0 to risk aversion. Let there be J possible outcomes in a lottery. Under EUT the probabilities for each outcome x_j, $p(x_j)$, are those that are induced by the experimenter, so expected utility is simply the probability weighted utility of each outcome in each lottery i:

$$EU_i = \sum_{j=1,J} [\, p(x_j) \times U(x_j) \,]. \tag{2}$$

The EU for each lottery pair is calculated for a candidate estimate of r, and the index

$$\nabla EU = EU_R - EU_L \tag{3}$$

calculated, where EU_L is the "left" lottery and EU_R is the "right" lottery as presented to subjects. This latent index, based on latent preferences, is then linked to observed choices using a standard cumulative normal distribution function $\Phi(\nabla EU)$. This "probit" function takes any argument between $\pm\infty$ and transforms it into a number between 0 and 1. Thus we have the probit link function,

$$\text{prob(choose lottery R)} = \Phi(\nabla EU) \tag{4}$$

Even though this "link function" is common in econometrics texts, it is worth noting explicitly and understanding. It forms the critical statistical link between observed binary choices, the latent structure generating the index ∇EU, and the probability of that index being observed. The index defined by (3) is linked to the observed choices by specifying that the R lottery is chosen when $\Phi(\nabla EU) > ½$, which is implied by (4).

Thus the likelihood of the observed responses, conditional on the EUT and CRRA specifications being true, depends on the estimates of r given the above statistical specification and the observed choices. The "statistical specification" here includes assuming some functional form for the cumulative density function (CDF). The conditional log-likelihood is then

$$\ln L(r; y, \mathbf{X}) = \sum_i [(\ln \Phi(\nabla EU)\times\mathbf{I}(y_i = 1)) + (\ln (1-\Phi(\nabla EU))\times\mathbf{I}(y_i = -1))] \tag{5}$$

where $\mathbf{I}(\cdot)$ is the indicator function, $y_i = 1(-1)$ denotes the choice of the right (left) lottery in risk aversion task i, and $\mathbf{X}$ is a vector of individual characteristics reflecting age, sex, race, and so on.

Harrison and Rutström [2008; Appendix F] review procedures that can be used to estimate structural models of this kind, as well as more complex non-EUT models. The goal is to illustrate how researches can write explicit maximum likelihood (ML) routines that are specific to different structural choice models. It is a simple matter to correct for multiple responses from the same subject ("clustering"), as needed.

It is also a simple matter to generalize this ML analysis to allow the core parameter r to be a linear function of observable characteristics of the individual or task. We would then extend the model to be $r = r_0 + R\times\mathbf{X}$, where r_0 is a fixed parameter and R is a vector of effects associated with each characteristic in the variable vector $\mathbf{X}$. In effect the unconditional model assumes $r = r_0$ and just estimates r_0. This extension significantly enhances the attraction of structural ML estimation, particularly for responses pooled over different subjects and treatments, since one can condition estimates on observable characteristics of the task or subject. In the present context we can introduce variables to reflect the answers to the RT questionnaires.

An important extension of the core model is to allow for subjects to make some *behavioral* errors. The notion of error is one that has already been encountered in the form of the statistical assumption that the probability of choosing a lottery is not 1 when the EU of that lottery exceeds the EU of the other lottery. This assumption is clear in the use of a non-degenerate link function between the latent index ∇EU and the probability of picking one or other lottery; in the case of the normal CDF, this link function is $\Phi(\nabla EU)$. If there were no errors from the perspective of EUT, this function would be a step function: zero for all values of $\nabla EU<0$, anywhere between 0 and 1 for $\nabla EU=0$, and 1 for all values of $\nabla EU>0$.

We employ the error specification originally due to Fechner and popularized by Hey and Orme [1994]. This error specification posits the latent index

$$\nabla EU = (EU_R - EU_L)/\mu \tag{3$'$}$$

instead of (3), where μ is a structural "noise parameter" used to allow some errors from the perspective of the deterministic EUT model. This is just one of several different types of error story that could be used, and Wilcox [2008] provides a masterful review of the implications of the alternatives. As $\mu \rightarrow 0$ this specification collapses to the deterministic choice EUT model, where the choice is strictly determined by the EU of the two lotteries; but as μ gets larger and larger the choice essentially becomes random. When $\mu=1$ this specification collapses to (3), where the probability of picking one lottery is given by the ratio of the EU of one lottery to the sum of the EU of both lotteries. Thus μ can be viewed as a parameter that flattens out the link functions as it gets larger.

An important contribution to the characterization of behavioral errors is the "contextual error" specification proposed by Wilcox [2011]. It is designed to allow robust inferences about the primitive "more stochastically risk averse than," and posits the latent index

$$\nabla EU = ((EU_R - EU_L)/\nu)/\mu \tag{3$''$}$$

instead of (3$'$), where ν is a new, normalizing term for each lottery pair L and R. The normalizing term ν is defined as the maximum utility over all prizes in this lottery pair minus the minimum utility over all prizes in this lottery pair. The value of ν varies, in principle, from lottery choice pair to lottery choice pair: hence it is said to be "contextual." For the Fechner specification, dividing by ν ensures that the *normalized* EU difference $[(EU_R - EU_L)/\nu]$ remains in the unit interval. The term ν does not need to be estimated in addition to the utility function parameters and the parameter for the behavioral error tern, since it is given by the data and the assumed values of those estimated parameters.

The specification employed here is the CRRA utility function from (1), the Fechner error specification using contextual utility from (3″), and the link function using the normal CDF from (4). The log-likelihood is then

$$\ln L(r, \mu; y, \mathbf{X}) = \sum_i [(\ln \Phi(\nabla EU) \times \mathbf{I}(y_i = 1)) + (\ln (1-\Phi(\nabla EU)) \times \mathbf{I}(y_i = -1))] \qquad (5'')$$

and the parameters to be estimated are r and μ given observed data on the binary choices y and the lottery parameters in **X**.

It is possible to consider more flexible utility functions than the CRRA specification in (1), but that is not essential for present purposes.

Once the utility function is estimated, it is a simple matter to evaluate the implications for risk aversion. Of course, the concept of risk aversion traditionally refers to "diminishing marginal utility," which is driven by the curvature of the utility function, which is in turn given by the second derivative of the utility function. Although somewhat loose, this can be viewed as characterizing individuals that are averse to mean-preserving increases in the variance of returns.

But there are also so-called "higher-order risk aversion" processes, known as prudence and temperance, that correspond to properties of the third and fourth derivative of the utility function, respectively (see Eeckhoudt and Schlesinger [2006]). Again loosely, these can be viewed as characterizing individuals that are averse to mean-preserving increases in the skewness and kurtosis of returns, respectively. The graphic on the next page summarizes these implications, again informally and under EUT.

For the CRRA utility function given by (1), and widely used for our estimates, the second derivative is $-rx^{-1-r}$ and the third derivative is $-(-1-r)rx^{-2-r}$. Hence it is a simple matter to evaluate if the individual exhibits prudence, for example.

B. Rank-Dependent Utility

The RDU model of Quiggin [1982] extends the EUT model by allowing for decision weights on lottery outcomes. The specification of the utility function is the same parametric specification (1′) and (1″) considered for EUT. To calculate decision weights under RDU one replaces expected utility defined by (2) with RDU

$$\mathrm{RDU}_i = \sum_{j=1,J} [\, w(p(M_j)) \times U(M_j) \,] = \sum_{j=1,J} [\, w_j \times U(M_j) \,] \tag{2′}$$

where

$$w_j = \omega(p_j + \ldots + p_J) - \omega(p_{j+1} + \ldots + p_J) \tag{6a}$$

for j=1,... , J-1, and

$$w_j = \omega(p_j) \tag{6b}$$

for j=J, with the subscript j ranking outcomes from worst to best, and $\omega(\cdot)$ is some probability weighting function.

We consider three popular probability weighting functions. The first is the simple "power" probability weighting function proposed by Quiggin [1982], with curvature parameter γ:

$$\omega(p) = p^{\gamma} \tag{7}$$

So $\gamma \neq 1$ is consistent with a deviation from the conventional EUT representation. Convexity of the probability weighting function is said to reflect "pessimism" and generates, if one assumes for simplicity a linear utility function, a risk premium since $\omega(p) < p \;\; \forall p$ and hence the "RDU EV" weighted by $\omega(p)$ instead of p has to be less than the EV weighted by p. The rest of the ML specification for the RDU model is identical to the specification for the EUT model, but with different parameters to estimate.

The second probability weighting function is the "inverse-S" function popularized by Tversky and Kahneman [1992]:

$$\omega(p) = p^{\gamma} \,/\, (\, p^{\gamma} + (1-p)^{\gamma} \,)^{1/\gamma} \tag{8}$$

This function exhibits inverse-S probability weighting (optimism for small p, and pessimism for large p) for $\gamma<1$, and S-shaped probability weighting (pessimism for small p, and optimism for large

p) for $\gamma>1$.

The third probability weighting function is a general functional form proposed by Prelec [1998] that exhibits considerable flexibility. This function is

$$\omega(p) = \exp\{-\eta(-\ln p)^{\varphi}\}, \tag{9}$$

and is defined for $0<p\leq 1$, $\eta>0$ and $\varphi>0$. When $\varphi=1$ this function collapses to the Power function $\omega(p) = p^{\eta}$. Of course, EUT assumes the identity function $\omega(p)=p$, which is the case when $\eta = \varphi = 1$. Many apply the Prelec [1998; Proposition 1, part (B)] function with constraint $0 < \varphi < 1$, which requires that the probability weighting function exhibit subproportionality (so-called "inverse-S" weighting). Contrary to received wisdom, many individuals exhibit estimated probability weighting functions that violate subproportionality, so we use the more general specification from Prelec [1998; Proposition 1, part (C)], only requiring $\varphi > 0$, and let the evidence determine if the estimated φ lies in the unit interval. This seemingly minor point often makes a major difference empirically.[8]

The construction of the log-likelihood for the RDU model with Power or Inverse-S probability weighting follows the same pattern as for EUT, with the parameters r, γ and μ to be estimated.

C. Dual Theory

The Dual Theory (DT) specification of Yaari [1987] is the special case of the RDU model in which the utility function is assumed to be linear. Hence diminishing marginal utility can have no influence on the risk premium, and the only thing that can explain the risk premium is "probability pessimism."

---

[8] One often finds applications of the one-parameter Prelec [1988] function, on the grounds that it is "flexible" and only uses one parameter. The additional flexibility over the Inverse-S probability weighting function is real, but minimal compared to the full two-parameter function.

D. Disappointment Aversion

Gul [1991] proposed a model of decision making under risk that allowed for a reference point for each lottery, and then posited that the decision maker might experience "disappointment" or "elation" relative to that reference point when evaluating the lottery.

Consider a lottery A with prizes x_i and objective probabilities p_i. Assume some utility function u(x), such as the CRRA function (1) proposed earlier, with parameter r. For a given value of r, we can then easily numerically evaluate the Certainty Equivalent (CE) of the lottery (in some special cases the CE is a closed-form expression, but in general it need not be).

Once the CE is calculated, we can define x+ to be the set of prizes greater than or equal to the CE, and x- to be the set of prizes worse than the CE. Then define the sum of probabilities for each of these components of the original lottery, for 4 possible outcomes:

$$p+ = \sum_{i=1,4} p_i \qquad \text{s.t. } x_i \in x+ \tag{10}$$

$$p- = \sum_{i=1,4} p_i \qquad \text{s.t. } x_i \in x- \tag{11}$$

We know that (p+) + (p-) = 1, by construction. Then we may construct a lottery based on A that reflects the prizes that are greater than the CE (A+) and a lottery based on A that reflects the prizes that are worse than the CE (A-) as follows:

$$A+: \quad \sum_{i=1,4} (p_i/p+)\, u(x_i) \qquad \text{s.t. } x_i \in x+ \tag{12}$$

$$A-: \quad \sum_{i=1,4} (p_i/p-)\, u(x_i) \qquad \text{s.t. } x_i \in x- \tag{13}$$

By construction, we know that A is now the lottery A+ with probability p+ and the lottery A- with probability (1-p+) = p-. To allow different weights for disappointment and elation define a function that weights these probabilities p+ and p- as follows:

$$\gamma(p+) = (p+) \,/\, [\, 1 + (1-(p+))\, \theta \,] \tag{14}$$

where $\theta \in (-1, \infty)$. The evaluation of the lottery is then just the disappointment-weighted evaluation of A+ and A-:

$$\gamma(p+) \times (A+) + [\,1 - \gamma(p+)] \times (A-) \tag{15}$$

When $\theta = 0$ we have $\gamma(p+) = p+$, and (15) is just the EUT evaluation of lottery A.

When $\theta > 0$ we place greater weight on A- than we would under EUT since $\gamma(p+) < p+$, and the decision-maker is said to be *disappointment averse*. When $\theta \in (-1, 0)$ we place greater weight on A+ than we would under EUT since $\gamma(p+) \geq p+$, and the decision-maker is said to be *elation loving*. The upshot for structural estimation is that we have two parameters, r and θ, to estimate.[9]

E. Cumulative Prospect Theory

The key innovation of CPT, in comparison to RDU, is to allow sign-dependent preferences, where risk attitudes depend on whether the individual is evaluating a gain or a loss. The concept of loss aversion, or sign-dependent preferences, is one that has been formalized in different ways in the literature. It is important to review the different formalizations, and implications for experimental design and estimation.

Kahneman and Tversky [1979] introduced the notion of sign-dependent preferences, stressing the role of the reference point when evaluating lotteries. They defined loss aversion as the notion that the disutility of losses weighs more heavily than the utility of comparable gains. Here is the key paragraph (p. 279) introducing the concept:

> A salient characteristic of attitudes to changes in welfare is that losses loom larger than gains. The aggravation that one experiences in losing a sum of money appears to be greater than the pleasure associated with gaining the same amount of money [...]. Indeed, most people find symmetric bets of the form (x, .50; -x, .50) distinctly unattractive. Moreover, the aversiveness of symmetric fair bets generally increases with the size of the stake. That is, if $x>y\geq 0$, then (y, .50; -y, .50) is preferred to (x, .50; -x, .50). According to [their] equation (1), therefore, $v(y)+v(-y)>v(x)+v(-x)$ and $v(-y)-v(-x)>v(x)-v(y)$. Setting $y=0$ yields $v(x)<-v(-x)$, and letting y approach x yields $v'(x)<v'(-x)$, provided v', the derivative of v, exists. Thus, the valuation function for

[9] We add some modest constraints on r so that the CE evaluations do not become degenerate.

losses is steeper than the value function for gains.

Note that at this stage there is no presumption that the difference between v(x) and -v(-x) be a constant, λ. Indeed, that assumption is never made in Kahneman and Tversky [1979], and appears later in the literature.

But when we say that the utility decrement of a unit loss, where the absolute value of (x-y) defines the unit here, is bigger than the utility increment of a unit gain, we need to be able to compare utility changes in the gain domain and the loss domain. This means that we cannot just have a utility scale that allows any order-preserving transformation: otherwise one could choose utility numbers such that the statement was true or false. This just means that we have to be more restrictive than allowing positive affine transformations, and restrict ourselves to defining utility on a ratio scale rather than an interval scale. This result is not easy to identify in the literature. Chateauneuf and Wakker [1999; Theorem 2.3, p. 142] present axiomatizations for CPT under objective risk that appear to allow value functions to be unique up to an *interval scale*, which is the same as allowing arbitrary *positive affine transformations*. But this is due to them imposing a "tradeoff consistency" assumption that effectively restricts the analysis to prospects that are either all defined over the gain domain or all defined over the loss domain. Thus one rules out so-called "mixed prospects," which are the general case and central to the robust empirical identification of utility loss aversion. The general version of this theorem, again applying solely for "loss prospects" and "gain prospects," is provided by Wakker and Tversky [1993; Theorem 4.3, p. 155], where the value function is again unique up to an *interval scale*. The general case, in fact referred to as "the truly mixed case" in the statement of the theorem, is provided by Wakker and Tversky [1993; Theorem 6.3, p. 159], and there the value function is only unique up to a *ratio scale*. In this case the value functions are unique up to transformations by a positive constant.

Note also the final discussion in the quote from Kahneman and Tversky [1979] about defining loss aversion in terms of the derivatives of the utility function around a zero reference point, which is y=0 in the quote. This suggestion anticipates later proposals for defining loss aversion from Köbberling and Wakker [2005] and others.

It is also worth noting that Kahneman and Tversky [1979] *assumed* that the decision weights for gains and losses were defined by the same probability weighting function (e.g., see Tversky and Kahneman [1992; p. 302] for an explicit statement of this assumption). This rules out "probabilistic loss aversion," to be defined later.

Tversky and Kahneman [1992; p. 309] popularized the functional forms we often see for loss aversion, using a CRRA specification of utility:

$$U(m) = m^{1-\alpha} / (1-\alpha) \text{ when } m \geq 0 \tag{16a}$$
$$U(m) = -\lambda[(-m)^{1-\beta} / (1-\beta)] \text{ when } m < 0, \tag{16b}$$

and where λ is the loss aversion parameter. Here we have the assumption that the degree of loss aversion for small unit changes is the same as the degree of loss aversion for large unit changes: the same λ applies locally to gains and losses of the same monetary magnitude around 0 as it does globally to any size gain or loss of the same magnitude. This is not a criticism, just a restrictive parametric turn in the specification compared to Kahneman and Tversky [1979].

Another way to write this, following Wakker [2010; p. 239] is as follows:

> The phenomenon can be modeled by a regular *basic utility function* u and a *loss aversion parameter* $\lambda > 0$, with $u(0) = 0$, and the utility function U of the form $U(\alpha) = u(\alpha)$ for $\alpha \geq 0$ [and] $U(\alpha) = \lambda u(\alpha)$ for $\alpha < 0$. The idea behind this definition is that u captures the intrinsic value of outcomes and satisfies usual regularity conditions such as being smooth and differentiable at $\alpha=0$, and λ is a factor separate from u. To distinguish U from u, we sometimes call U the *overall utility*. The unqualified term utility will continue to refer to U. We use the following scaling conventions for u and λ, which are plausible if u is differentiable at 0 and is approximately linear on the small interval [-1, 1]. [...].
>
> $$u(1) = 1,\ u(-1) = -1,\ U(1) = 1, \text{ so that } \lambda = -U(-1)$$
>
> This scaling convention was implicitly adopted by Tversky and Kahneman [1992], who chose $u(\alpha) = \alpha^{\theta}$ for gains and $u(\alpha) = \alpha^{\theta'}$ for $\alpha < 0$, with a θ' that possibly differs from θ. These scaling choices amount to the convention of $u(1) = 1$ and $u(-1) = -1$, [...].

To anticipate, and remove a technical side issue, the analytical problems for loss aversion come when the coefficients θ and θ′ differ. Using the definition proposed by Köbberling and Wakker [2005], loss aversion is infinite if $\theta > \theta'$ and zero if $\theta < \theta'$. So we shall assume they are the same, which is one of the assumptions Köbberling and Wakker [2005; §7] and Wakker [2010; §9.6] propose themselves.[10] Extending the notation from Wakker [2010; p.239] to be explicit, and using the specification (16a) and (16b) from Tversky and Kahneman [1992; p. 309] with $\alpha=\beta$, we then have

$$U(m) = u(m) = m^{1-\alpha} / (1-\alpha) \text{ for } m \geq 0 \tag{17a}$$
$$U(m) = -\lambda\, u(-m) = -\lambda[(-m)^{1-\alpha} / (1-\alpha)] \text{ for } m < 0. \tag{17b}$$

Does this discussion have any implications for what choice tasks can be used to identify loss aversion? If one has choice data *solely in the gain domain* it is possible to estimate the basic utility function, as defined above. Then one can look at choices defined *solely in the loss domain*, and estimate the λ that best explains them, in effect "holding basic utility constant" based on the choices in the gain domain. This is one solution to the identification problem of picking a λ and a basic utility function at the same time. Note that one could use choices from the gain domain and choices from the mixed domain just as well, *or* choices from the loss domain and choices from a mixed domain, but one needs to have *some* choices from the gain or loss domain. In our design we have no problems of this kind, since we have gain-frame, loss-frame *and* mixed-frame lotteries.

---

[10] The same issue with identification arises if one employs more flexible functional forms than the power utility function or CRRA functions in general. For example, the two-parameter Expo-Power utility function popularized by Holt and Laury [2002] collapses to CRRA for certain parameter values, so it also runs into the same theoretical issues. It is curious to see how the literature on CPT tries it "kill off" power utility *per se*, rather than accept the restriction that is implied for identification. For example, Schmidt and Zank [2008; p. 214] note that "The preceding analysis shows that strong risk aversion and CPT exclude power utility. More generally, the [loss aversion ratio] is not well defined for power utility under CPT unless the two powers are equal." These are two different statements. The first says that one can never use power utility, the second says that power utility is valid under a simple restriction on the powers in the gain and loss frame. Later, they note (p.214) that Köbberling and Wakker [2005; §7] "... showed that power utility is problematic for the index of loss aversion. They suggested an alternative parametric form..." Again, the word "problematic" just means that one should not try to estimate *separate* powers in the gain and loss frame, not that one has to discard power utility, CRRA utility, or its natural generalizations such as Expo-Power.

Does one have to have mixed-frame lotteries to be able to estimate the loss aversion parameter λ? It is apparent that λ does not affect preferences over two gain-frame lotteries, or for that matter between two loss-frame lotteries. In the former case λ literally plays no role, and in the latter case it scales the total utility of all lotteries equally, so cannot change their *ranking*. But changes in λ does change the *numerical value* of the difference in the *total* utility between two loss-frame lotteries. If we have to constrain the *intrinsic* utility of gains and losses to be the same functional form, for reasons discussed above, then the fact that λ changes the numerical value of the difference in *total* utility of two loss-frame lotteries does provide a formal basis for estimating λ. As λ varies, the difference in *total* utility of loss-frame lotteries changes, hence the probability of the observed choice over loss-frame lotteries changes, hence the likelihood of all observed choices changes, and hence there is a λ, *ceteris paribus* all other parameters in the model, that maximizes the likelihood of observing all choices. However, even if one can theoretically identify λ with just data from the gain and loss frame, the use of mixed frame choices makes that identification practically easier. It is appropriate to think of "the identification problem" as a matter of degree, even though the expression is often used in an all-or-nothing sense.[11]

Apart from the critical role of the same intrinsic utility function, one must also *avoid* applying the "contextual utility" normalization of Wilcox [2011] that was appropriate for EUT and RDU, and affected (3″). This is why it is important to be clear about the scaling restrictions on intrinsic utility

[11] Bruhin, Fehr-Duda and Epper [2010; p. 1382] claim that one cannot identify λ in the all-or-nothing sense by just comparing preferences over lotteries in the loss domain (one of their "lotteries" is a certainty-equivalent, but that is not essential to the argument). This is correct, and noted above: preference over two loss-frame lotteries are not affected by the value of λ. But this is precisely why one needs to estimate λ jointly with lottery choices from the loss frame *and* the gain frame. So we disagree with their conclusion that "... when there are no mixed lotteries available, estimating such a parameter is neither feasible nor meaningful." We would agree that in the absence of mixed lotteries it is not *easy* to estimate λ reliably, so there is an identification problem in the broader sense if one does not include mixed lotteries in the design.

clarified earlier, that intrinsic utility be unique up to a *ratio* scale (and not the weaker interval scale).[12]

There is a clear statement of the "exchange rate assumptions" needed to define loss aversion in Abdellaoui, Bleichrodt and Paraschiv [2007; p.1662], as well as a tabulation of the range of definitions that have been proposed in the literature. For instance, Fishburn and Kochenberger [1979] and Pennings and Smidts [2003] defined loss aversion as $U'(-x)/U'(x)$, Tversky and Kahneman [1992] as $-U(-1)/U(1)$, Bleichrodt, Pinto and Wakker [2001] as $-U(-x)/U(x)$, and Schmidt and Traub [2002; p.235] as $U(x)-U(y) \leq U(-y)-U(-x)\ \forall\ x>y\geq 0$. One can make the exchange rate assumptions formally *de minimus* by defining an index of loss aversion solely in terms of the directional derivatives at the reference point, $U'_{\rightarrow}(0)/U'_{\leftarrow}(0)$, as proposed by Köbberling and Wakker [2005] and Booij and van de Kuilen [2009]. But this has the very unfortunate effect, as honestly emphasized by Wakker [2010; p. 247], that *global* properties of loss aversion are being driven by very, very *local* properties of estimated utility functionals,[13] and that puts a great strain on empirics and functional form assumptions.

For comparability with the "base camp" of the mountain of estimations of CPT, we follow Tversky and Kahneman [1992] and define utility loss aversion as $\lambda \equiv -U(-1)/U(1)$. Hence the empirical strategy is to evaluate estimates of α and β, and then infer λ by evaluating the implied utility function at ± 1. Estimates of all three parameters are then used, along with estimated decision

[12] It is common to assume that u(x) is weakly increasing in x, and that $u(0) = 0$, hence that $U(0)=0$. This implies that $\lambda > 0$, so that the utility values for losses are all negative. If we do not impose the arbitrary normalization $u(0)=0$, then weak monotonicity only implies that $\lambda \geq -1$, since we then have $U(x) \geq U(-x)$ if we have the same "intrinsic" utility function u(x). Values of λ that are negative have the same interpretation in this case as values of λ that are positive. The general, non-parametric definition of loss aversion proposed by Schmidt and Traub [2002; p. 235], that $U(x) - U(y) \leq U(-y) - U(-x)\ \forall\ x > y \geq 0$, evaluates identically whether λ is positive or negative, providing monotonicity holds. We set $u(0) = [u(\tau) + u(-\tau)]/2$ for $\tau \rightarrow 0$, and the exact, small value of τ makes no difference to results; we use $\tau = \$0.50$. This assigns a value to u(0) that depends on the specific utility function and parameters being estimated, and ensures that we have monotonicity at 0 if $\lambda<0$. For comparability with other studies we always report $|\lambda|$.

[13] In other words, the utility loss aversion for a loss of one penny is the same proportionally as the utility loss aversion of one million dollars.

weights, to evaluate each lottery using (10a) and (10b).

What if the probability weighting functions for the gain domain differ from the probability weighting functions for the loss domain? There is nothing *a priori* in CPT to rule this out, and good reasons to want to de-couple the extent of probability weighting in the gain and loss frames. Even if the basic utility functions for gains and losses are linear, and conventional loss aversion is absent ($\lambda=1$), differences in the *decision weights* for gains and losses could induce the same behavior as if there were utility loss aversion. This is called "probabilistic loss aversion" by Schmidt and Zank [2008; p.213]. Imagine that there is no probability weighting on the gain domain, so the decision weights are the objective probabilities, but that there is some probability weighting on the loss domain. Then one could easily have losses weighted more than gains, from the implied decision weights.

To see the point intuitively, assume a power probability weighting function, so statements about concave or convex probability weighting apply for all objective probabilities. Then one simply needs to have the probability weighting function for losses be convex (overweighting) and the probability weighting function for gains be linear for there to be probabilistic loss aversion.[14] In this case we have probability neutrality for gains and probability pessimism for losses, implying, *ceteris paribus*, risk neutrality over gains and risk aversion over losses. These assumptions are stronger than needed, but illustrate the importance for estimates of the "utility loss aversion" parameter λ of

[14] Note that the effect on risk aversion of overweighting and underweighting for losses is the opposite of the effect for gains. As noted by Neilson and Stowe [2002; p.34] CPT "weights extreme outcomes first," the *best* outcome under gains and the *worst* outcome under losses. Appendix B discusses this point in more detail. We generally avoid the expressions "optimism" and "pessimism" in CPT, since it can lead to confusion in general. Wakker [2010; p. 289] offers an excellent definition of these terms, which generalizes from RDU to CPT: optimism (pessimism) is when an improvement in the rank is associated with a higher (lower) decision weight. Of course, to add potential semantic confusion, this definition further assumes the use of what Wakker [2010; §7.6] calls gain-ranks rather than loss-ranks. The former is now the default, but the latter was used in early studies on RDU, such as the classic by Chew, Karni and Safra [1987], since it is more natural when referring to cumulative density functions in statistics.

allowing flexible degrees of probability weighting in the gain and loss domains.[15]

We allow flexibility in the probability weighting for losses and gains with the power probability weighting function by using

$$\omega(p) = p^{\gamma+} \text{ for } m \geq 0 \tag{18a}$$
$$\omega(p) = p^{\gamma-} \text{ for } m < 0 \tag{18b}$$

and where the p in question is the objective probability associated with that specific m. For the inverse-S function we use

$$\omega(p) = p^{\gamma+} \,/\, (\, p^{\gamma+} + (1\text{-}p)^{\gamma+} \,)^{1/\gamma+} \text{ for } m \geq 0 \tag{19a}$$
$$\omega(p) = p^{\gamma-} \,/\, (\, p^{\gamma-} + (1\text{-}p)^{\gamma-} \,)^{1/\gamma-} \text{ for } m < 0. \tag{19b}$$

For the Prelec function we use

$$\omega(p) = \exp\{\, -\eta^{+} \, (-\ln p)^{\varphi+} \,\} \tag{20a}$$
$$\omega(p) = \exp\{\, -\eta^{-} \, (-\ln p)^{\varphi-} \,\}. \tag{20b}$$

The construction of the log-likelihood for the CPT model follows the same pattern as for EUT and RDU, with the parameters α, λ, γ+ and γ- (or $-\eta^{+}$, φ^{+}, η^{-} and φ^{-}) as well as a Fechner error term μ, to be estimated. One difference is that the "contextual utility" normalization is inappropriate on theoretical grounds, as noted earlier. A second difference is that loss frame lotteries have their decision weights based on a rank-ordering from worst prize to best, rather than the rank-ordering from best to worst used for gain frame lotteries (and all lotteries under RDU). Appendix B explains this difference. A final difference is that *mixed* frame lotteries are "parsed" into a gain frame version and a loss frame version, which are then separately evaluated. The evaluation of the overall mixed frame lottery is then the *sum* of these two components, as explained in detail in §2 of Appendix B.

[15] There can be a probabilistic loss aversion or loss seeking effect even if the probability weighting functions for gains and losses are the same. The point is rather that if sign-dependence is the key insight of CPT then one should not *a priori* hardwire one of the determinants of (probabilistic) loss aversion with that constraint.

3. Estimates

We first estimate each of the alternative models with data pooled across all subjects, to allow a characterization of representative individual behavior. We initially assume homogeneous preferences to keep things simple, illustrate results, and explain how we classify behavior as best characterizing by one model. We then allow for heterogeneous preferences by estimating for each individual.

The bottom line of our analysis is a comparison of the performance of the five models, shown in Figure 5 for the sample of 177 undergraduates and in Figure 6 for the sample of 95 MBA students with higher stakes. We conclude that RDU is the best characterization overall, followed by EUT. There is virtually no support for DT, DA or CPT.

The simple reason for the poor performance of CPT is that our subjects locally asset integrate over the frames presented to them. As stressed earlier, this is a key difference between CPT and the other models.[16] We expect DT and EUT to underperform RDU, since they are each nested in RDU. Finally, the poor showing for DA is consistent with extensive evidence from the early experimental work on EUT rejecting the Betweenness Axiom (BWA)[17] that it uses as an alternative to the Independence Axiom (IA): see Camerer and Ho [1994; p. 191] and Starmer [2000; p.358].

---

[16] There is good reason to be verbose on this point, since a number of studies referenced in Section 5 estimate an RDU model but casually refer to it as CPT. By this they presumably mean that the RDU model contains the same probability weighting as CPT, but that is only on the gain frame. RDU and CPT are otherwise very different models.

[17] The difference is easy to explain, and important to understand for the positive evidence in favor of RDU. The usual IA states that preferences over lotteries A and B are not changed if we consider some lottery consisting of a p chance of A and a (1-p) chance of C and some lottery consisting of a p chance of B and a (1-p) chance of C, for any C and all p. In words, preferences over two lotteries are not affected by adding a common consequence C with the same probability weight. The BWA simply restricts C to be some mixture of A or B. The important consequence of this change from the IA to the BWA is that indifference curves within the Marschak-Machina probability simplex are *still linear* but do *not have to be parallel*, as required under EUT. The BWA axiom also underlies the Chew-Dekel class of risk preferences, which play a critical role in Epstein-Zinn preferences in finance.

A. Estimates for the Representative Agent

Although individual-level estimates are our focus, it is useful to consider the estimates for each model when it is assumed to characterize every subject. Table 1 presents maximum likelihood estimates for each model. EUT shows moderate risk aversion, at a level generally consistent with many years of evidence from laboratory experiments. We can also see the EUT estimate as a "descriptive" way to flag that there is a risk premium for the representative agent, even if the other models decompose it differently.

With DT and the Prelec probability weighting function, allowing $\eta > 0$ and $\varphi > 0$, we have evidence of probability pessimism for almost all the range of probabilities, as shown in Figure 7. The left panel of Figure 7 shows the estimated probability weighting function, and the right panel shows the implied decision weights for the ranked outcomes, using an equi-probably reference lottery to see the pure effect of rank-dependent probability weighting. The top (second) [bottom] line in the right panel shows a two-outcome lottery in which both outcomes have a probability of ½ (⅓) [¼]. So we see in each case that the worst outcome is given greater decision weight than the best outcome, but the best outcome is given slightly more weight than the intermediate outcomes for lotteries with 3 or 4 outcomes. This probability weighting function has the popular "inverse-S" shape, but is predominately convex. Hence DT generates a risk premium, consistent with the implication from the EUT model.

Turning to the RDU model, we focus here just on the specification with the Prelec probability weighting function. We now get a very different pattern of probability weighting than the DT model, as demonstrated by Figure 8 and comparison to Figure 7. This probability weighting function is concave, implying risk-loving behavior *ceteris paribus* the effect of the curvature of the utility function. Since we know from the EUT and DT estimates that there is a risk premium overall, we can infer that there must be a more concave utility function than EUT in order for the net effect of probability optimism and diminishing marginal utility to generate a modest risk premium. This is exactly what we find, in Table 1, with the CRRA coefficient estimate of 0.70 compared to the EUT estimate of 0.45. For both DT and RDU we can easily reject the hypothesis that there is no probability weighting (i.e., that $\eta = \varphi = 1$).

The DA model estimates in Table 1 imply statistically significant concave utility ($r = 0.70 > 0$) and disappointment aversion ($\theta = 0.51 > 0$). This model has to do as well as EUT, since EUT is nested, when $\theta = 0$. It might be expected to do as well as RDU, since if there are only two prizes then the DA model literally collapses to RDU (Abdellaoui and Bleichrodt [2007]); on the other hand, we have up to four prizes per lottery in these data, and in those cases RDU and DA are not the same.

Finally, we consider the CPT model with the Prelec probability weighting function. Table 1 presents the estimates, and Figures 9 and 10 visualize their implications. We find concave utility in the gain frame ($\alpha > 0$), linear utility in the loss frame ($\beta \approx 0$), and mild evidence for utility loss aversion ($\lambda > 1$). The top left panel of Figure 9 shows the estimated "intrinsic" utility functions, and the top right panel then shows the "full" utility functions. These full utility functions are the same in the gain frame as the intrinsic utility function, but the full utility function in the loss frame also incorporates the effect of utility loss aversion, and is shown in the solid line.

The CPT estimates for probability weighting imply the pattern shown in Figure 10: classic "inverse-S" probability weighting for gains, and concave utility weighting for losses. The implication of concave probability weighting for losses is to put greater (lower) weight on the worst (best) outcomes, so that we would have probabilistic loss aversion if there had been no probability weighting for gains.

The "hit rates" of successful predictions for EUT, DT, RDU, DA and CPT are 71%, 70%, 73%, 72% and 76%, respectively, for the undergraduates using house money to cover any losses.[18] The log-likelihoods are -10818, -10952, -10456, -10722 and -10936. So by both metrics, RDU dominates EUT which in turn dominates CPT. Of course, CPT and {EUT, RDU} are not nested.[19]

For the non-nested model comparisons we use tests that compare models by looking at the likelihoods for the same observation, and defining a statistics on those observation-specific comparisons, rather than the sum of likelihoods. The two most popular tests are the Vuoung test and the Clarke test, described in Harrison and Rutström [2009]. Using them we are able to test the null hypothesis that the two models are equally close to the true specification, and that one cannot discriminate between them. Each test allows us to say which model if "favored," but also provides some statistical confidence in the rejection of the null in the direction of the favored model. Using these tests we can draw a strong conclusion for the representative agent comparisons: the CPT model is favored over each of the EUT, DT, RDU and DA models in terms of both tests, each test allows one to reject the null hypothesis of non-discrimination with p-values below 0.01, and the distribution of data underlying the test statistics is non-Gaussian so the Clarke test should be used.[20] As it happens both tests lead to the same conclusion: CPT wins.

Both tests generate the same conclusions, which flows primarily from the huge sample size with all of the pooled choices. There is some evidence that the ratio of log-likelihoods is non-Gaussian, and leptokurtic, when sample sizes are less than 500: the asympotic convergence to a Gaussian distribution of the ratio of log-likelihoods is slow for smaller sample sizes, and the distribution tends to resemble a double exponential (Clarke and Signorino [2010; p. 377]). In that case the Clarke test is much more reliable. We will see below that this point is of some importance

[18] Unless otherwise stated, we only report results for the most flexible probability weighting specification for DT, RDU and CPT, the Prelec function.

[19] If one uses the "information criteria" AIC and BIC, which allow comparisons of non-nested models using *ad hoc* punishment terms for additional parameters, the same ranking applies.

[20] The only exception was a *p*-value between the RDU and CPT models of 0.04.

when evaluating individual subjects, since the sample size is 100 there and we do see differences between the conclusions from the two tests.

A final way to evaluate the risk preferences of the representative agent is by means of a mixture model, following Harrison and Rutström [2009]. Focusing only on the mixture between the EUT and CPT models, and the RDU and CPT models,

The first thing to see is that the RDU model characterizes 66% of the choices: indeed, even the EUT model characterized 68% of the choices in a comparable mixture model.

The second thing to see is that we have a very concave utility function for the RDU choices, but an optimistic probability weighting function, more or less offsetting each other. In fact, if we estimate at EUT and CPT model, we find that the utility function is concave, *suggesting* that the net effect of the offsetting processes in the RDU model is to be risk averse. Here is the probability weighting function under RDU:

The third thing to see is that we have a concave utility function for gains under CPT, but a linear utility function for losses. Utility loss aversion is significant. We also find very little probability weighting in either gain or loss frames, as shown below. This differs from some prior estimates in which we get very little utility loss aversion, but significant probabilistic loss aversion, at least at the pooled level.

B. Hypothesis Tests to Discriminate Between Models

When typing individuals as EUT, DT or RDU we have the benefit of a direct hypothesis test that $\omega(p) = p$. Similarly, when typing individuals as EUT or DA we can similarly directly test the hypothesis that $\theta = 0$. But when we have CPT we have to either estimate a mixture model for each subject or apply some non-nested hypothesis tests. The former can be challenging numerically for samples of just 100 observations, and the latter can be reliably undertaken if we have log-likelihoods for each observation and for each model being preferred.[21] Hence we first consider how the nested hypotheses tests allow us to discriminate between the EUT, DT, RDU and DA models, and then consider the extension to discriminating between these models and CPT.

Figure 11 displays the models selected using the non-nested hypothesis tests with respect to EUT. The left panel displays the distribution of *p*-values on these tests, with one *p*-value for each subject. We select the winning model for each subject based on log-likelihoods, and then check if the null hypothesis of EUT can be rejected at the 1%, 5% or 10% significance level. Unless the alternative to EUT exhibits statistically significant rejection of EUT, we retain the EUT classification.[22] We find a relatively low fraction of EUT-consistent subjects for samples from this population, just over 25%. There are virtually no subjects classified as DT, and very few classified as DA. The modal risk preference type is RDU, with the Prelec probability weighting function dominating.

Now consider extending this model discrimination exercise to include CPT. Start with a subject for whom RDU is favored over CPT, such as subject #2 (in turn, EUT was rejected for this subject in the comparison with RDU). In Figure 12 we show the "jagged" distribution of the log-likelihood ratios that form the basis of the Vuoung test of the non-nested model. We also show a fitted Normal distribution to this empirical distribution, since the Vuoung test requires that this distribution be Normal, as it is asymptotically.[23] It is apparent that the empirical distribution is more peaked than the normal, consistent with the tendency in samples below 500 for this distribution to be leptokurtic. In fact, in the last subtitle line we show the *p*-value from the Chen and Shapiro [1995]

[21] When sample sizes are large enough, the mixture model is preferred for reasons spelled out in Harrison and Rutström [2009; §5].

[22] We also check the DA and RDU models using non-nested hypothesis tests, and they do not change the rankings based solely on the aggregate log-likelihoods.

[23] In fact, the requirement is that the distribution be a *unit* Normal, and the fitted Normal shown here, and implicitly used in the statistical tests of Normalcy, is not constrained to have zero mean or standard deviation of 1, but that only makes our tests conservative with respect to rejecting Normalcy.

test of Normalcy, as implemented by Brzezinski [2012]; this test has been shown to be more robust than the popular Shapiro-Wilk test. This *p*-value is less than 0.01, implying that we can reject the hypothesis that the log-likelihood ratios are distributed Normally, and hence we must use the Clarke test rather than the Vuoung test.[24]

Doing so, we find that the Clarke test favors the RDU model over the CPT model, and that the *p*-value on the null hypothesis of non-discrimination between the RDU and CPT models is below 0.01. Hence, at the 5% significance level we can reject that null, and infer that the RDU model is favored over the CPT model for this subject. The hit rates of the fitted models are also consistent with this conclusion: the CPT hit rate is only 71% compared to the RDU hit rate of 80%. In this instance we see the importance of determining if the conditions for the Vuoung test are met: the *p*-value on non-discrimination for that test is only 0.49, and we would not have concluded that the RDU model was superior to the CPT model.

An example in which CPT wins is subject #7, shown in Figure 13. In this instance both non-nested hypothesis tests favor the CPT model over RDU, as well as CPT have a superior hit rate. Again, the Chen-Shapiro test rejects the assumption required for the Vuoung test, although in this instance that is moot.

There are many instances in which EUT dominates CPT, and Figure 14 illustrates the case of subject #8. In this instance the Clarke test favors EUT, and indeed the hit rate for EUT is greater than the hit rate for CPT. However, the *p*-value on the Clarke test being unable to statistically discriminate between EUT and CPT is 0.09, so in this instance the subject is classified as EUT since

[24] One can adjust the Clarke and Voung test statistics to "punish" models with relatively more parameters (e.g., Clarke and Signorino [2010; p. 376]), but the correction factors are the same *ad hoc* ones noted earlier for the AIC and BIC "information criteria," and again have no convincing methodological foundation. These corrections would only strengthen a conclusion *not* to adopt the CPT model, so they would not increase the share of individuals classified as CPT decision-makers.

CPT is not favored in a statistically significant manner.

A similar example arises for subject #3, shown in Figure 15, where the CPT model is favored over the EUT, but not in a statistically significant manner. Again, we therefore classify this subject as an EUT decision-maker, in the absence of statistically significant evidence to the contrary.

Turning from individual instances, we can characterize the general trend of these hypothesis tests. Figure 16 shows the fraction of cases in black for which the "base model" named on the far left is favored by the non-nested hypothesis test indicated by the Chen-Shapiro test of Normalcy of the log-likelihood ratios. The fraction of cases in light blue show where the CPT model is favored compared to the base model. Of course, each subject is classified as one or other of these two base models, but these unconditional results illustrate the pattern more clearly than just focusing on shares for the preferred base model for each subject.

In general we see that the CPT model fares better when compared to the EUT model rather than the RDU model, as one might expect from the added flexibility of the RDU model. But these shares, again, do not tell us whether the favored model was favored *in a statistically significant manner*. For that we need to look to the *p*-values on the null hypothesis of non-discrimination, as illustrated for the specific examples considered earlier. Figure 17 shows these for the cases in which the non-nested hypothesis test favored the CPT model (i.e., the sub-sample from Figure 16 shown in light blue). These are the cases we care about to understand why the CPT model, although favored, was not favored in a statistically significant manner. Figure 17 shows the *average p*-values in this case, although one could eaqually look at the median or interquartile range. We show one of the 9 complete distributions of *p*-values later, to verify this point. Figure 17 suggests that it does not matter which non-nested hypothesis test, Vuoung or Clarke, is used when comparing RDU and CPT models: in both cases the average of *p*-values is well above 10%. Hence, for the subjects classified as RDU from the nested hypothesis test, which is around 60% of the sample (see Figures 5 and 6), it is very unlikely that the non-nested hypothesis test would *positively and statistically significantly* favor the CPT model (and recall that this is the average for the cases in which the CPT model was favored).

However, Figure 17 does suggest that it matters whether the Vuoung or Clarke test is used when comparing EUT and CPT models, or DA and CPT models. The Vuoung test has average *p*-values that are between 0.10 and 0.19, suggesting that there could be a decent fraction below 1%, 5% or 10%. But the Clarke test has very high *p*-values, virtually ensuring that the CPT model would *not be positively and statistically significantly* favored over the EUT or DA model. The final piece in this story, then, is to see how often the Vuoung test is used when the CPT model is favored. Figure 18 makes it clear, if one considers the tiny values on the bottom axis, that the Clarke test is almost always used. This figure shows the average *p*-values of the Chen-Shapiro test of Normalcy of the ratio of log-likelihoods: rejecting that hypotheses implies that we must use the Clarke test instead of the Vuoung test, as noted above for the 4 individual examples.

Finally, Figure 19 shows a complete distribution of *p*-values for the Clarke non-nested hypothesis comparing RDU Prelec and CPT Prelec, in contrast to the average shown in Figure 17 in the light blue bars. The kernel density approximation does "bleed" below 0 and 1 slightly, but the point is clear: the vast bulk of the *p*-values are greater than the 1%, 5% and 10% levels shown in red, dashed lines.

4. Previous Literature

Surely we are not the first to estimate a structural version of CPT? As it happens, we are not, but it is rather remarkable to see how light the previous evidence is when one weights the experimental and econometric procedures carefully. Moreover, a recent trend seems to be to declare any evidence for probability weighting, even if only in the gain domain, as evidence for CPT when it is literally evidence for RDU. Table 2 summarizes our review of the literature we are aware of, focusing only on controlled experiments, which has been the original basis of empirical claims for CPT.

Tversky and Kahneman [1992] gave their 25 subjects a total of 64 choices. Their subjects received $25 to participate in the experiment, but rewards were not salient, so their choices had no monetary consequences. The majority of data from their experiments used an elicitation procedure that we would now call a multiple price list, in the spirit of Holt and Laury [2002]. Subjects were told the expected value of the risky lottery, and 7 certain amounts were presented in a logarithmic scale, with values spanning the extreme payouts of the risky lottery. The subject made 7 binary choices between the given risky lottery and the series of certain amounts. To generate more refined choices, the subject was given a second series of 7 certainty equivalents for the same risky lottery, zeroing in on the interval selected in the first stage. This variant is called an *iterative* multiple price list by Andersen, Harrison, Lau and Rutström [2006]. Furthermore, "switching" was ruled out, with the computer program enforcing a single switch between the risky lottery and the certain values. This variant is called a *sequential* multiple price list by Andersen, Harrison, Lau and Rutström [2006]. All risky prospects used two prizes, and there were 56 prospects evaluated in this manner. One half of these prospects were in the gain frame, and one half were in the loss frame, with the latter being a "reflection" of the former in terms of the values employed.

A further 8 tasks involved mixed-frame gambles. In these choices the subject was asked to

Fill-In-the-Blank (FIB) by entering a value \$x that would make the risky lottery (\$a, ½; \$b, ½) equivalent to (\$c, ½; \$x, ½), for given values of a, b and c. The probabilities for the initial 56 choices over gain frame or loss frame choices were 0.01, 0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95 and 0.01, whereas the sole probability for the 8 mixed-frame choices was ½.[25]

Tversky and Kahneman [1992] estimate a structural model of CPT using non-linear least squares, and at the level of the individual. Remarkably, they then report the *median* point estimate, for each structural parameter, over the 25 estimated values. So over all 25 subjects, and using our notation, the median value for α was 0.88, the median value of λ was 2.22, the median value of γ+ was 0.61, and the median value of γ- was 0.69.[26]

These parameter estimates are remarkable in three respects, given the prominence they have received in the literature. First, whenever one sees point estimates estimated for individuals, one can be certain that there are many "wild" estimates from an *a priori* perspective, so reporting the median value alone might be quite unrepresentative of the average value, and provides no information whatsoever on the variability across subjects. Second, there is no mention at all of standard errors, so we have no way of knowing, for example, if the oft-repeated value of λ is statistically significantly different from 1. Third, the median value of any given parameter is not linked in any manner to the median value of any other parameter: these are *not the values of some representative, median subject*, which is often how they are implicitly portrayed.[27] The subject that actually generated the median value of λ,

[25] Wakker [2010; p. 175] sharply admonishes anyone that only uses one probability to elicit risk attitudes. Of course, Tversky and Kahneman [1992] used several probabilities in the gain frame and in the loss frame, so it is surprising that they did not do likewise in the mixed frame. No obvious "all-or-nothing" identification problems arise from their choice set design overall, but identification of probabilistic loss aversion is surely improved, in the broader sense, if one allows various probabilities in mixed frame lotteries.

[26] They also estimated β and apparently obtained *exactly* the same median value as α, which is quite remarkable from a numerical perspective.

[27] Tversky and Kahneman [1992; p. 312] do note that the "parameters estimated from the median data were essentially the same." It is not clear how to interpret this sentence. It may mean that the median certainty-equivalents for the initial 56 choices, and the median values of \$x for the final 8 choices, were combined to form a synthetic "median subject," and then estimates obtained from those data. The expression

for instance, might have had any value for α, β, γ+ and γ-.

These shortcomings of the Tversky and Kahneman [1992] study have not, to our knowledge, led anyone to replicate their experiments with salient rewards and report complete sets of parameter estimates with standard errors. The fault is not that of Tversky and Kahneman [1992], who otherwise employed quite modern methods, but the subsequent CPT literature. Anybody casually using these estimates as statistically representative must not care about rigor in empirical work.

Camerer and Ho [1994] is a remarkable study, with many insights. It was also one of the first to propose and estimate a structural model of CPT using maximum likelihood (§6.1). The data employed were choice patterns from a wide range of studies, but the analysis was explicitly restricted to the gain frame (p. 188). Hence it could be said to be the first structural estimation of the RDU model, but not of a CPT model including losses.

Wu and Gonzalez [1996] focus entirely on the probability weighting function. They stress the point that they estimate the probability weighting function without having to make assumptions about utility functions, and view the need to make those assumptions as a methodological flaw. The reason it is said to be a flaw is that inferences about the probability weighting function could be confounded by mis-specifications of the true utility function (p.1678). They propose a simple method for eliciting probability weights based on a series of choices with only two common outcomes, \$200 or \$240. Hence one could normalize utilities of these outcomes to 0 and 1, and avoid making any further assumptions about the utility function. Unfortunately this procedure was implemented in a non-salient, hypothetical choice task, and only for the gain frame (§4). When Wu and Gonzalez [1996] undertake maximum likelihood estimation, via a non-linear least squares method, they assume a power utility function and also restrict themselves to gain frame choices (§5).

"median data" does not lead one to suspect that it was any one actual subject. Nor is there any reference to standard errors for these estimates.

One could adapt the Wu and Gonzalez [1996] method for eliciting a probability weighting function for the gain frame to eliciting functions for the gain *and* loss frame, but they did not do so. Gonzalez and Wu [1999] estimate (non-parametric) probability weighting functions *and* utility functions for 10 subjects based on elicited certainty-equivalents for two-outcome lotteries solely in the gain frame. They at least employed salient rewards for their small number of subjects.

Harbaugh, Krause and Vesterlund [2002] paid for one of the 24 lotteries studied. Each lottery had two outcomes, with zero payment possible in every lottery. In half the lotteries the second payment was positive, and the other half of lotteries had a negative second payment; thus, there were no mixed frame lotteries. Each decision was between one of the lotteries and a certain amount, which was usually the expected value of the lottery. Decisions were presented to subjects on separate plastic cards, with each lottery presented as a pie chart with a "spinner" in the middle of the circle. Extra care was given to the method of task presentation, since subjects were as young as five years old. They do not undertake structural estimation of the CPT model, claiming (p.83) that, "Given our data it is not possible to simultaneously estimate both the probability weighting function and the value function." They do not consider utility loss aversion at all.

Mason, Shogren, Settle and List [2005] evaluate behavior over risky lotteries defined solely in a loss frame. They do not consider gain frame choices or mixed frame choices, but they do employ salient, real rewards.

Stott [2006] examines a wide range of parametric functional forms for CPT, but only considers data from hypothetical tasks defined over the gain frame.[28]

Fehr-Duda, Gennaro and Schubert [2006] paid subjects for one of 50 binary choices over lotteries with two outcomes. Half of the battery of losses were for gains, half were for losses, and

[28] Stott [2006; p.113] notes that one choice was incentivized by scaling prizes down from nominal amounts up to £40,000 to an actual payment amounts up to £5. Average salient payments were just £2.13. We view this as effectively hypothetical.

there were no mixed frame choices. For each lottery, an ordered MPL with 20 certain amounts was used to elicit a certainty equivalent. The certain amounts spanned the two outcomes of the lottery, so each subject faced 50 MPLs each with 20 rows. The utility loss aversion parameter λ was not estimated, because of the absence of mixed frame lotteries (p. 295).

Fennema and van Assen [1998], Abdellaoui [2000], Etchart-Vincent [2004], Schunk and Betsch [2006], Abdellaoui, Bleichrodt and l'Haridon [2008] and Booij and van de Kuilen [2009] are widely cited as having used the "tradeoff method" to estimate the utility function for losses. Fennema and van Assen [1998], Etchart-Vincent [2004] and Booij and van de Kuilen [2009] used hypothetical survey questions, with no real consequences. Abdellaoui [2000; p. 1502] and Schunk and Betsch [2006; p. 389] used real incentives for the gain frame, but hypothetical survey questions for the loss frame; neither asked any questions in the mixed frame. Abdellaoui, Bleichrodt and l'Haridon [2008] also asked real questions in the gain frame, but only hypothetical survey questions in the loss and mixed frames. Brooks and Zank [2005] used real losses, and focused on testing certain implications for choice patterns from utility loss aversion, not estimating the full CPT structure. In a similar vein, Brooks, Peters and Zank [2014] used real losses from a house endowment, and generated choice predictions based on assumed parametric values for a standard CPT specification. No CPT model was estimated from the 105 binary choices each subject made over gain, mixed and loss frames.

Rieskamp [2008] uses "slightly real" rewards and all three frames. Subjects made binary choices over lotteries with outcomes between +€100 and -€100, one of 180 choices was selected for payment and realization, and then 5% of the outcome added or subtracted from an endowment of €15. So the rewards were salient, but not substantial. Nonetheless, this is a great advance from virtually all other studies. The structural estimates employed both α and β in power utility functions, with no discussion of the implications for identifying utility loss aversion. As it happened, the estimates of these two parameters were virtually identical, as in Tversky and Kahneman [1992]. The utility loss aversion parameter was constrained to be greater than 1, ruling out utility loss seeking. And the parameters for the Inverse-S probability weighting functions were constrained to be less than 1 for both gains and losses. Pooled over all subjects, the estimates (p. 1455) were $\alpha = \beta = 0.91$, $\lambda = 1$, $\gamma+ = 0.69$ and $\gamma- = 0.71$. It is an open question what these estimates would be if λ had not "hit" the imposed lower boundary value.

Boij, van Praag and van de Kuilen [2010] estimate parametric models of CPT, but use hypothetical survey questions.

Bruhin, Fehr-Duda and Epper [2010] estimated parametric models of CPT that assumed that the utility loss aversion parameter λ was 1, noting wryly that "our specification of the value function seems to lack a prominent feature of prospect theory, loss aversion..." (p. 1382). They did this because their design only included lotteries in the gain frame and the loss frame, and none in the mixed frame. Estimation of utility loss aversion is logically impossible without mixed frame choices. They did provide real incentives for decisions, and employed an endowment of house money just as we did.

Pachur, Hanoch and Gummerum [2010] studied inmates in a UK prison, as well as UK non-prisoners. Choices were hypothetical, as the inmates received no compensation of any kind, and the non-prisoners received only a fixed £3 pound participation payment that was non-salient.

Nilsson, Rieskamp and Wagenmakers [2011] utilized the same "slightly real" data of Rieskamp [2008], but applied a Bayesian hierarchical model to estimate structural CPT parameters. They recognized the identification problem with power utility specifications when $\alpha \neq \beta$ indirectly. They initially simulated data using the popular point estimates from Tversky and Kahneman [1992], to test the ability of their model to recover them. They found that their model underestimated λ and that α was estimated to be much lower than β, rather than $\alpha \approx \beta$. They concluded (p.89) as follows:

> It is likely that these results are caused by a peculiarity of CPT, that is, its ability to account for loss aversion in multiple ways. The most obvious way for CPT to account for loss aversion is by parameter λ (after all, the purpose of λ is to measure loss aversion). A second way, however, is to decrease the marginal utility at a faster pace for gains than for losses. This occurs when α is smaller than β. Based on this reasoning, we hypothesized that the parameter estimation routines compensate for the underestimation of λ by assigning lower values to α than to β; in this way, CPT accounts for the existing loss aversion indirectly in a manner that we had not anticipated.

Of course, this is just the theoretical identification noted earlier, and discussed in Köbberling and Wakker [2005; §7] and Wakker [2010; §9.6]. In any event, they optionally estimate all models with $\alpha = \beta$, and avoid this identification problem. Using the Inverse-S probability weighting function they reported Bayesian posterior modes (standard deviations) over the pooled sample of $\alpha = \beta = 0.91$ (0.16), $\lambda = 1.02$ (0.26), $\gamma+ = 0.68$ (0.11) and $\gamma- = 0.89$ (0.19). Unlike Rieskamp [2008], they did not constrain λ to be greater than 1. These estimates are the Bayesian counterparts of random coefficients: hence each parameter is a distribution, which can be summarized in several ways. Reporting the mode is a more robust alternative to the mean, given the symmetric nature of their visual display of estimates, and the standard deviation provides information on the estimated variability across the 30 subjects, each making 180 binary choices. They find no evidence for utility loss aversion. Figure 21 shows the two probability weighting functions estimated, and implied decision weights. There is *very* slight evidence of probabilistic loss aversion for small probabilities, since there is slight risk loving over gains and extremely slight risk aversion for losses. For large probabilities this evidence suggests probabilistic loss seeking, albeit modest.[29]

[29] They also report (Table 2, p.91) ML estimates for each of the 30 subjects, and comment about the relative imprecision of these estimates compared to those obtained from the pooled Bayesian hierarchical methods. We agree with this likely outcome from individual-level estimates, as noted earlier, even when there are 180 binary choices per subject. Earlier they anticipated this finding, noting (p. 87) that they "... illustrate how single-subject maximum likelihood, one of the most popular estimation methods for CPT (e.g. Harrison & Rutström 2009; Harless & Camerer, 1994; Stott, 2006), can produce extreme, implausible point estimates for parameters estimated with high uncertainty." The first two studies references here did not in fact estimate at the level of the individual, as claimed, and Stott [2006] used hypothetical choice data.

Glöckner and Pachur [2012] undertook incentivized experiments, presented subjects with 138 binary choices over two-outcome lotteries spanning the gain, loss and mixed frame. A house endowment of €22 was used to cover potential losses of up to €9.90, from one lottery choices that was selected to play out.[30] Structural CPT estimates were generated, and one of their metrics for selecting parameters reflected likelihoods, rather than the unweighted hit rate. However, it appears that their estimation procedures do not generate standard errors, as illustrated by the tests of the hypothesis of stability of choices over two sessions.[31] Median estimates of parameters across individuals are reported (Table 4, p.27), following the unfortunate procedure of Tversky and Kahneman [1992], so one cannot say what any individual or representative agent's parameters were. EUT is compared (p. 29), but only with respect to the unweighted hit rate; there is no comparison to RDU, although a long list of *ad hoc* heuristics (Table 2, p. 26) are compared in terms of unweighted hit rates.

von Gaudecker, van Soest and Wengström [2011] estimated parametric models of CPT that assumed a complete absence of probability weighting, on both gain and loss frames. They note clearly (p.675) that their specification entails

> ...departures from the original prospect theory specification. [...] it does not involve nonlinear probability weighting because our goal is to estimate individual-level parameters, and the dimension of the estimation problem is large already. Adding a parameter that is highly collinear with utility curvature in our experimental setup would result in an infeasibly large number of parameters, given the structure of our data. Furthermore, typical probability weighting functionals develop the highest impact at extreme probabilities, which are absent from our experiment.

---

[30] An unfortunate, but popular, use of a "lab currency" allowed them to state outcomes ranging between -€1000 and +€1200. These amounts were scaled down by 100 if chosen for payment. This procedure is unattractive, since it only affects behavior if subjects exhibit money illusion and are unable to infer the true payoff in the natural currency. If subjects exhibit money illusion then there is a loss of control over stimuli, by definition, since one does not know how the illusion manifests itself (e.g., non-linearly). We prefer to deal with the budgetary consequences of presenting monetary amounts in the natural currency.

[31] They consider correlations of parameter estimates for each subject between the two sessions (p. 28), rather than a direct test of the hypothesis that the estimate *distributions* are the same.

Unfortunately these justifications are tenuous. The fact that the goal is individual-level estimation does not, by itself, have any theoretical implications for why one can pick and choose aspects of the CPT model. Indeed, adding one or two parameters for probability weighting, assuming one of the popular one-parameter specifications and the possibility of constraining probability weighting to be the same in the gain and loss frames, does add to the dimensionality of the estimation problem. But numerical convenience is hardly an acceptable rationale for mis-specification of the CPT model. Colinearity with utility curvature is actually a theoretical point of some importance, and to be expected, and not an econometric nuisance. Indeed, it extends to colinearity with the utility loss aversion parameter, unless one assumes away *a priori* the possibility of probabilistic loss aversion. If one parameter plays a significant role in explaining the risk premium for an individual, then assuming it away surely biases conclusions about the strength and even sign of other psychological pathways. The final point, about not having sufficient variability in probabilities to estimate probability weighting functions, is even less clear. Their initial lottery choices varied the probability of the high prize from 0.25 to 0.5, 0.75 and 1; then their second stage choice interpolated the probability weights between one of these gaps (0 to 0.25, 0.25 to 0.5, 0.5 to 0.75, or 0.75 to 1) in grids of roughly 10 percentage points. Even from the first stage choices, if one assumes the popular Power or Inverse-S function then one only needs one interior probability to allow estimation. In fact, they always have the three interior probabilities of the first stage, and typically have refinements within one of those intervals. In sum, these arguments sound as though they were constructed "after the fact" of extensive numerical and econometric experimentation, and in the face of *a priori* unreliable numerical results.

von Gaudecker, van Soest and Wengström [2011] employed a design in which all payments were to be sent to participants 3 months after their choices were made. This was to allow the design to vary the time of resolution of risk (now or in the future), without confounding that treatment with the timing of payment and discount rates. Their payoff configurations (Table 1, p. 669) include gain frame lotteries, mixed-frame lotteries, and no loss frame lotteries. Four of the seven payoff configurations have all risk resolved at the time of choice, although by means of a computer realization (raising issues of credibility).

Zeisberger, Vrecko and Langer [2012] estimate a structural CPT model from experimental data from 89 students, who earned €60 in an experiment a month prior, with payment only for the two sessions. One in ten students were paid, based on their choices for one random task out of 30. They elicited CE for lotteries in the gain, mixed and loss frames, using the Becker, De Groot and Marschak [1964] procedure. They estimated a "full" model for each subject in which all CPT parameters are jointly estimated using maximum likelihood methods. For some reason standard errors needed to be generated by bootstrapping (e.g., Table 5, p. 375), and no hypothesis tests of parameters are presented. Median estimates are presented (Table 4, p. 373), but at least interquartile ranges are also presented. No estimates for a representative agent are presented. Individual point estimates are presented (Table 5, p. 375ff.), and exhibit some "wild" estimates. This may be due to the small number of choices for each subject, although if the CE is reliably elicited it embeds more information than a binary choice. No comparison between CPT and other models is presented.

Abdelloui, l'Haridon and Paraschiv [2013] estimated parametric models of an RDU model defined over gains, but referred to this as a CPT model even if there were no losses at all in the stimuli. They did use real incentives, and told 65 couples that "they could be selected to play out one of their choices for real..."; it is not clear if one of the 65 would be selected for salient rewards, or this means that there was some probability that each couple could be selected. In any event, this is not a CPT model since losses played no part.

In summary, Table 2 shows that very few studies that we are aware of use real, salient incentives for gain, loss and mixed frames. Those that meet these methodological criteria are shaded.

5. Limitations and Extensions

We are well aware of limitations of our results. On the other hand, we defend them as the appropriate place to start a rigorous examination of the general empirical validity of the CPT model, claims about the importance of loss aversion in general, and the psychological pathway for loss aversion. We consider limitations from the perspective of theory, experimental procedure, and econometrics, recognizing that these are not independent domains.

A. Theoretical Issues

The first theoretical point is the vexing question of the specification of the "right" reference point. Kahneman and Tversky [1979] were explicitly agnostic on this issue, Tversky and Kahneman [1992] were silent, and the issue has been dormant until the recent development of "endogenous reference point" models by Kőszegi and Rabin [2006][2007] and Schmidt, Starmer and Sugden [2008]. Obviously any theoretical specification of the reference point that differs from the framed reference point in our experiments will make a difference to the effect of frames, since that reference point acts to define what choices fall into which frame.[32] These theoretical specifications are, however, surprisingly vague as to how they are to be operationalized, and their rigorous evaluation remains open pending those specifications. Of course, the CE in Disappointment Aversion models provides one, early endogenous reference point specification.

A second theoretical point is global asset integration, by which we mean the assumed manner in which earnings within the laboratory are combining with extra-lab income or wealth. In

[32] Without developing a theory of "the" reference point, Harrison and Rutström [2008; p.95-98] evaluate a wide range of parametrically assumed reference points, and construct a "profile likelihood" for each of them. The reference point with the best profile likelihood was not $0, as assumed here, but some positive amount possibly reflecting some "homegrown reference point" that the subject brought to the lab based on expected earnings. What is relevant here is not their method of finding the empirically best-performing reference point, which was a-theoretical, but that the structural parameter estimates for utility loss aversion were much more in accord with *a priori* beliefs when that alternative reference point was assumed.

one sense the issue of global asset integration, which raises the calibration critique of estimates of risk aversion from small stakes, is one reason one might want to rigorously model loss aversion. Rabin [2000; p.1288] used loss aversion as the primary throw-away explanation of why one actually observes subjects picking safer lotteries over riskier lotteries, even when they perfectly integrate "wealth" with income from experimental lotteries:

> What *does* explain risk aversion over modest stakes? While this paper provides a "proof by calibration" that expected-utility theory does not help explain some risk attitudes, there are of course more direct tests showing that alternative models better capture risk attitudes. [...] Many of these models seem to provide a more plausible account of modest-scale risk attitudes... [...] indeed, what is empirically the most firmly established feature of risk, loss aversion, is a departure from expected-utility theory that provides a direct explanation for modest-scale risk aversion.

We disagree with many, in fact all, of the assertions here, but the point is that non-EUT specifications are viewed as one way of accounting for calibration puzzles.[33] Moreover, the historical evidence for CPT has been accumulated with utility functions defined solely over experimental income.

B. Experimental Procedures

<u>Earned Endowments</u>

If someone had undertaken a prior task, with some real task, and with real earnings rather than some artefactual earnings from "house money," would they exhibit greater loss aversion?

This is an easy extension to make to our design. Following Laury, McInnes and Swarthout [2009], we modified our procedures to provide subjects with a quiz of 15 questions in which they

---

[33] These calibration puzzles were independently developed by Hansson [1988] and Rabin [2000], and rest on an empirical premiss that subjects exhibit risk aversion over a wide enough range of wealth and lotteries defined over small stakes. Building on an ingenious design independently due to Cox and Sadiraj [2008; p.33] and Wilcox [2013], Harrison, Lau, Ross and Swarthout [2016] show that this premiss is strikingly false for subjects drawn from the same population as the experiments reported here. It is *not* false for other populations of interest, such as adult Danes, at least for the finite range of wealth considered in the experiments.

could earn \$80 or \$40 depending on their knowledge of current events, American history, and geography. We also explained that any earnings or losses from later choice tasks would be added to or subtracted from the earnings from the quiz. We therefore generated an endowment that applied equally to all three frames: the "house money" of our main experiments only applied to the mixed frame and loss frame. It would be artificial, and disingenuous, to generate an "earned endowment" that can only be retained if someone was running the risk of losing it. Appendix D contains the introductory text to the experiment that explained the connection between the quiz that generated an earned endowment and the later choice tasks. It also contains the quiz itself, which was designed to be relatively easy to score 8 or more correct answers, and hence \$80. The main instructions in Appendix C were modified to explain that any gains or losses would be added or subtracted from the quiz earnings, but were otherwise the same as those in the main experiments.

The choice tasks were those given to undergraduates, involving maximal losses of \$70, so nobody would lose their entire endowment if they had earned \$80. All net earnings were on top of the show-up fee. We had prepared an alternative set of choice questions with maximal losses of \$40 in the event that someone failed to earn \$80 from the quiz, but that did not happen.

Over two sessions, 58 undergraduate GSU subjects completed these experiments, which are obviously not cheap to run because of the earned endowment. Figure 22 shows the pooled estimates for this samples of undergraduates, to be compared to Figure 9 for undergraduates with house money to cover losses. There is an increase in the estimate of β from 0.06 to 0.23 as we move from house money to earned endowments, and λ increases from 1.34 to 1.76. The form of probability weighting over gains and losses is virtually identical.

At the level of individual estimation, however, this treatment did increase the fraction of subjects classified as CPT, at the expense of those classified EUT and RDU. Figure 20 shows the classifications, to be contrasted with Figure 5.

Somebody wanting to defend CPT might argue that our earned endowment task was facile, and that it amounted in effect to just another "house money" treatment in the minds of subjects. Alternative procedures for generating earned endowments, with time to integrate them into extra-lab wealth, have been proposed (e.g., Bosch-Domènech and Silvestre [2010] and Cárdenas, De Roux, Jaramillo and Martinez [2014]) and could be evaluated. Clearly, however, at some point the burden has to rest on advocates of CPT to propose an operationally meaningful way to endow subjects and then evaluate behavior in an econometrically rigorous manner.

Alternative Elicitation Methods

Would there be any effect from using alternative elicitation methods than binary choices? One popular alternative is to elicit a certainty-equivalent of some lottery, allowing one to directly infer the risk premium conditional on believing that the certainty-equivalent has been reliably elicited. The use of the open-ended, "fill in the blank," Becker, De Groot and Marschak [1964] elicitation method for certainty-equivalents is controversial: many experimenters believe that it performs poorly in practice, for various reasons. One can devise iterative multiple price lists, in the sense of Tversky and Kahneman [1992] and Andersen, Harrison, Lau and Rutström [2006], that can "drill down" in a series of ordered, binary choice tasks to effectively elicit a tight interval for the certainty-equivalent. The incentive-compatibility of these methods are more likely to be understood by subjects than the open-ended methods.

C. Econometric Issues

We characterize heterogenous preferences by estimating at the level of the individual, with a design that allows that because there are 100 binary choices for each individual. Another way to account for unobserved individual heterogeneity is to estimate structural models using random coefficients that reflect the latent population distribution of the parameters across subjects. Econometric methods for the estimation of non-linear systems, of the kind we encounter with the EUT, DT, RDU, DA and CPT structural models, have been developed by Andersen, Harrison, Hole, Lau and Rutström [2012]. Although we see these as valuable techniques to characterize heterogeneity, we do not expect them to fundamentally alter our conclusions about the ability of CPT to explain the broad pattern of observed behavior.

Finally, there is a perennial issue in econometrics of parametric specifications versus non-parametric specifications. In fact, there are three issues here: the use of parametric assumptions about the utility and probability weighting functions, the use of non-parametric predictions of theories (typically about choice patterns), and the use of parametric assumptions about the stochastic error processes. In each case we are open to the use of non-parametrics, but caution that there is a tradeoff in power when one does so, of course conditional on us having "good" or "flexible" parametric specifications.

6. Conclusions

Cumulative Prospect Theory does not obviously dominate alternative specifications of decision making under risk. In all treatment, RDU explains the behavior of more subjects, although with earned endowments the superiority of RDU is not as great as with the use of house money. The reason for this poor performnce of CPT, where "poor" is in relation to the hyerbole found in the behavioral literature, is that sSubjects locally asset integrate over frames, and then apply probability weighting consistent with an RDU model. Of course, extended versions of CPT might explain these behavior, and should be evaluated in future research, but the core CPT model does not fare well.

Another theme of our results is to discourage the sloppy habit of *defining* the CPT model in terms of the qualitative properties of specific parameter values. For example, some behave as if the CPT *model* claims that "individuals overweight low probabilities and underweight high probabilities," "probability weighting in the loss domain is the same as probability weighting in the gain domain," or that "loss aversion drives risk premia," when these just happen to be specific *instances* of the model.[34] This is a semantic matter, but an important one. None of these claims emerge from our experiments and econometric analysis. A more serious version of this problem of definitions is referring to estimates of an RDU model as "prospect theory."

Some defenders of the CPT model claim, correctly, that the CPT model exists "because the data says it should." In other words, the CPT model was born, in Kahneman and Tversky [1979], from a wide range of stylized facts culled from parts of the cognitive psychology literature. If one is to take the CPT model seriously and rigorously then it needs to do a much better job of explaining the data than we see here.

---

[34] These claims are familiar from the literature, but one can document one for completeness. Wakker [2010; p.234] makes the remarkable empirical claim that "I think that more than half of the risk aversion empirically observed has nothing to do with utility curvature or with probability weighting. Instead, it is generated by loss aversion..."

Figure 1: Illustrative Display of Lottery Choices

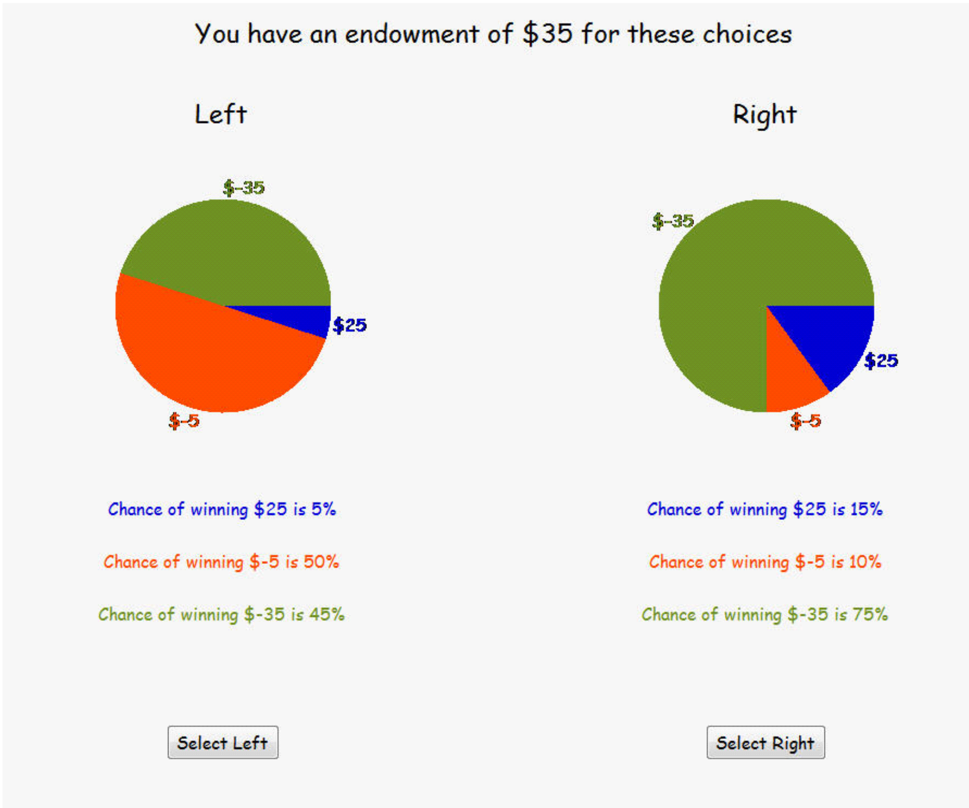


Figure 2: Marschak-Machina Triangles for Lotteries Used with Undergraduates

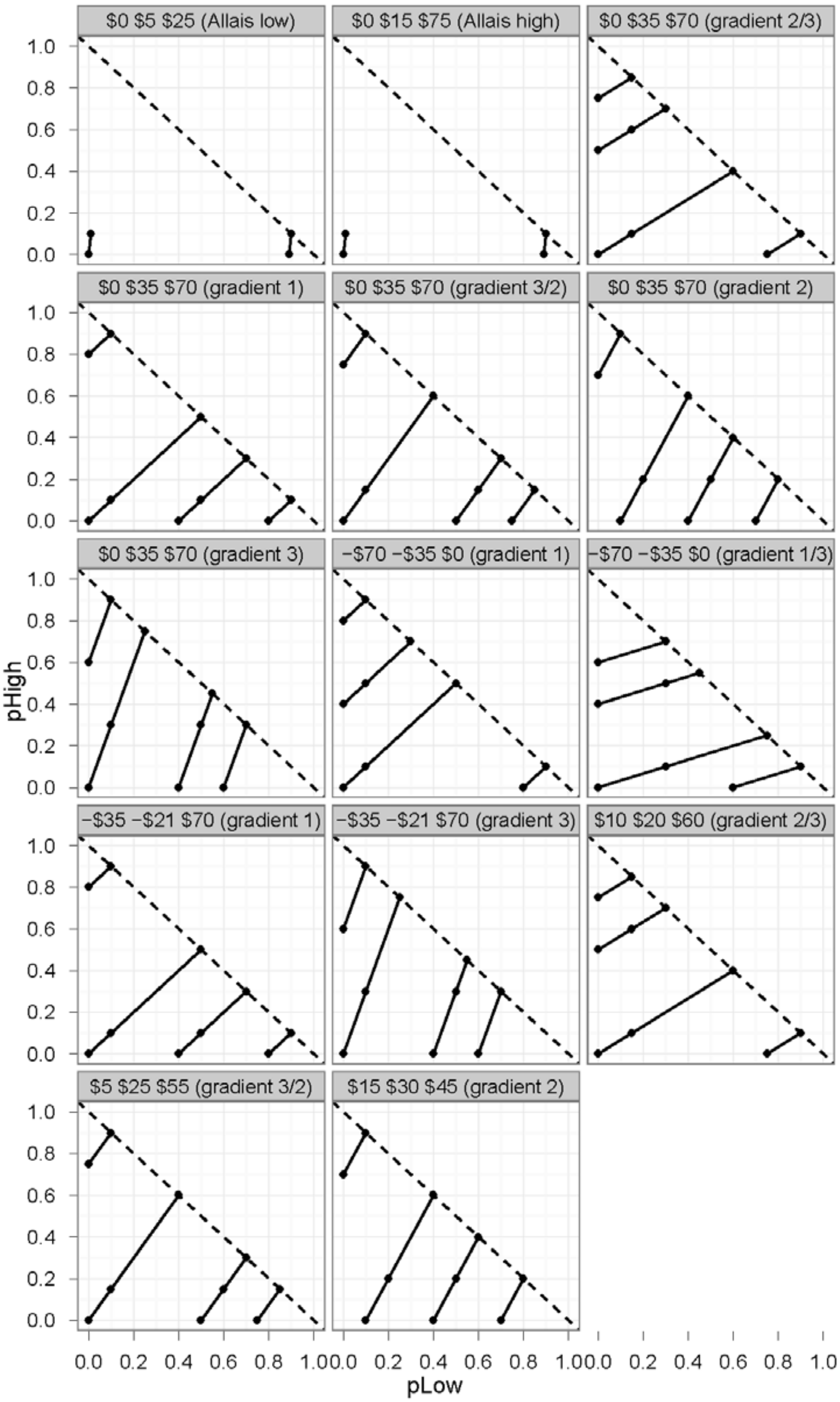


Figure 3: Marschak-Machina Triangles for Lotteries Used with MBA Students

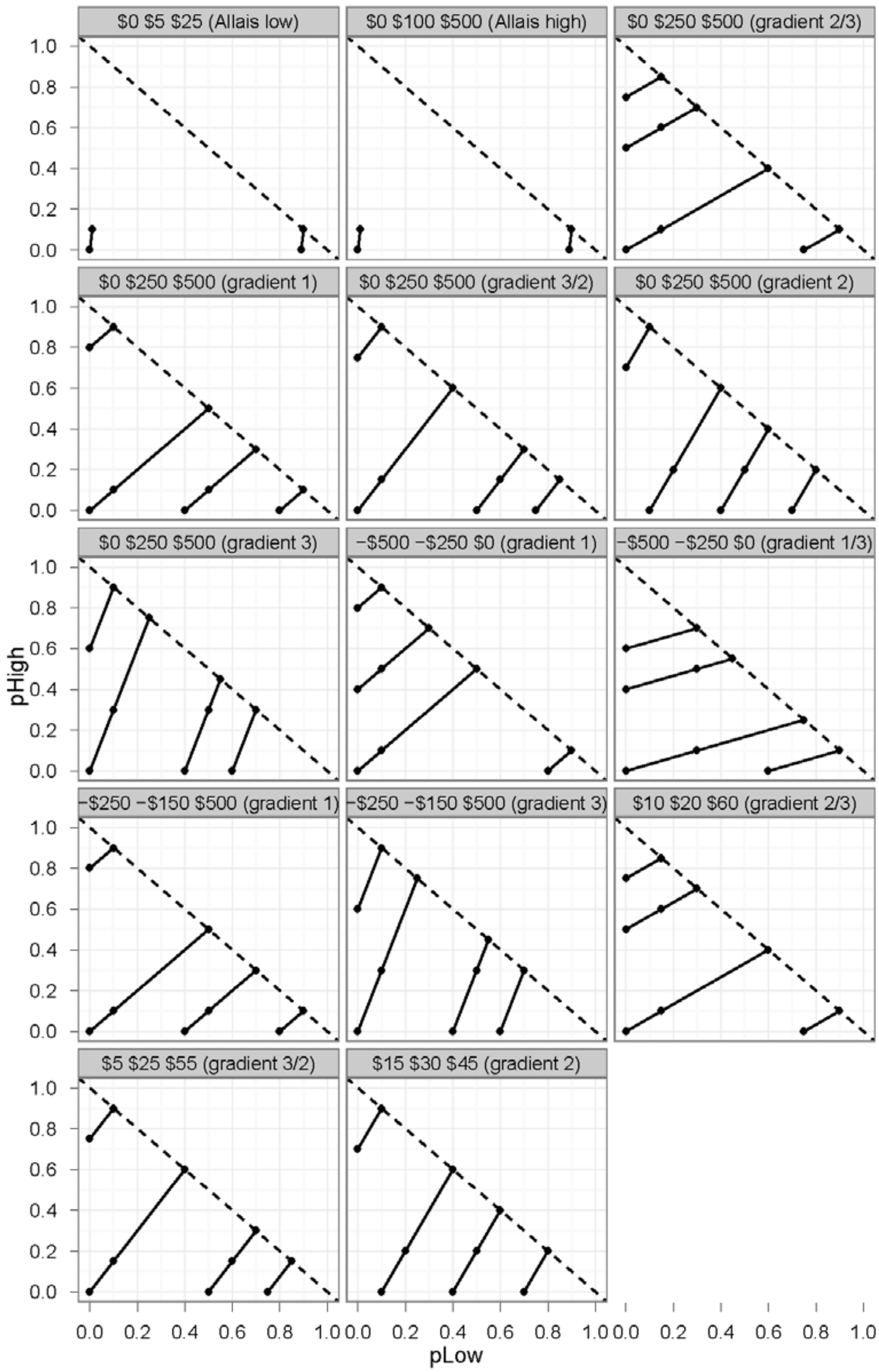


Figure 4: Unconditional Marshak-Machina Display of Lotteries

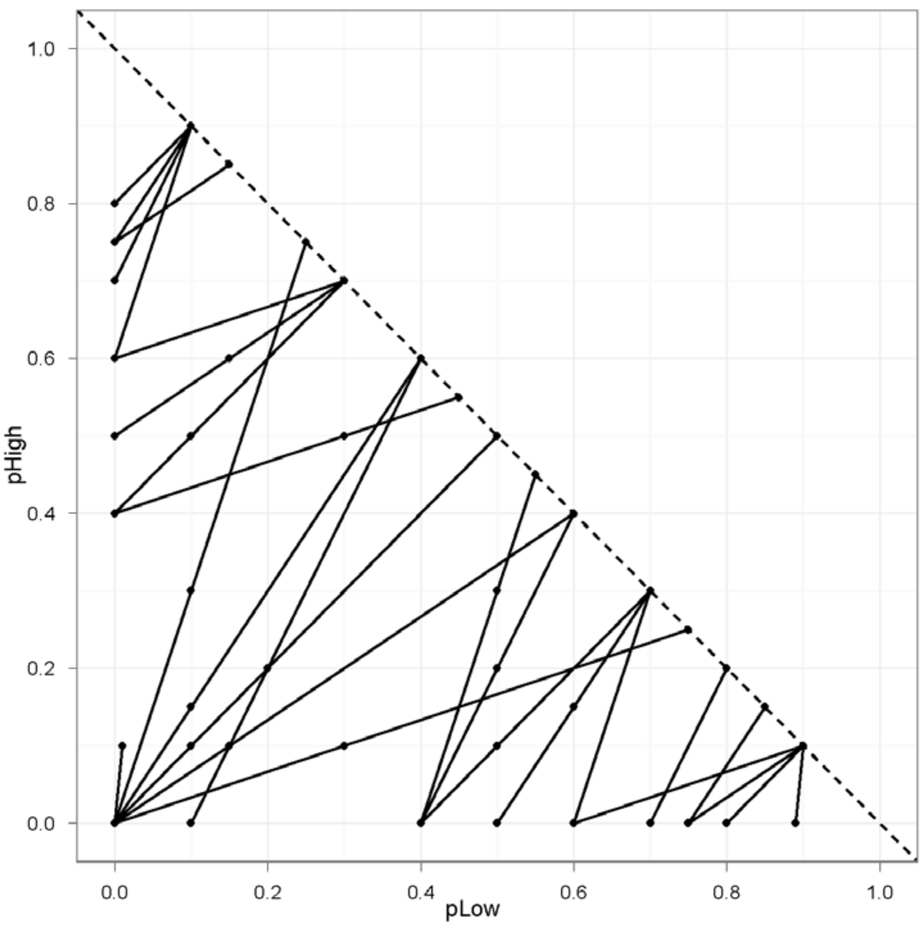


Figure 5: Classifying Subjects as EUT, DT, RDU, DA or CPT

Undergraduate students, endowed with house money to cover losses
N=177, one *p*-value per individual and a 5% signficance level
Estimates for each individual of EUT, DT, RDU, DA and CPT specifications

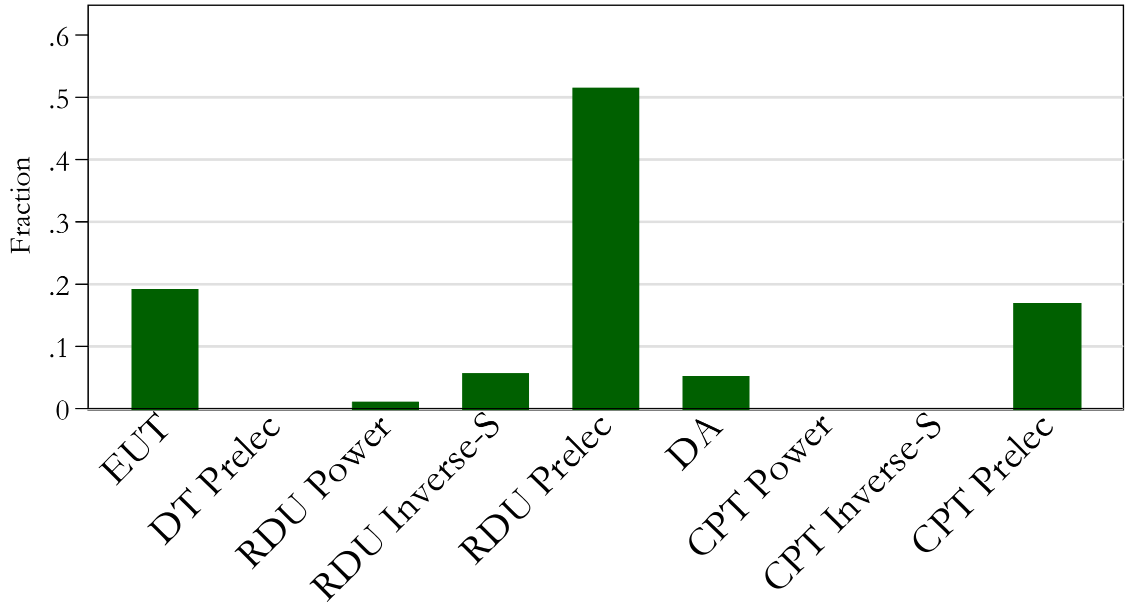


Figure 6: Classifying Subjects as EUT, DT, RDU, DA or CPT

MBA students, endowed with house money to cover losses
N=94, one *p*-value per individual and a 5% signficance level
Estimates for each individual of EUT, DT, RDU, DA and CPT specifications

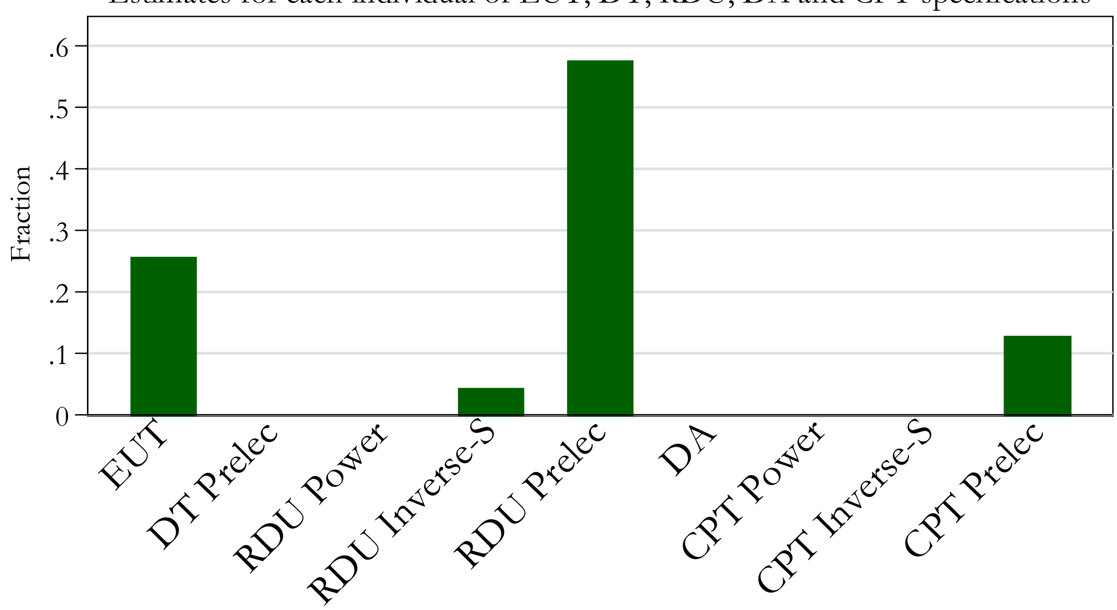


Table 1: Estimates for EUT, DT, RDU, DA and CPT Models with Pooled Data

N=177 undergraduates with house money to cover losses
DT, RDU and CPT estimates with the Prelec Probability Weighting Function

Parameter	Point Estimate	Standard Error	*p*-value	95% Confidence Interval	
			A. Expected Utility		
r	0.45	0.028	<0.001	0.39	0.50
			B. Dual Theory		
η	1.20	0.041	<0.001	1.12	1.28
φ	0.55	0.026	<0.001	0.50	0.61
			C. Rank Dependent Utility		
r	0.70	0.019	<0.001	0.66	0.74
η	0.51	0.025	<0.001	0.46	0.56
φ	0.18	0.010	<0.001	0.84	0.99
			D. Disappointment Aversion		
r	0.68	0.018	<0.001	0.65	0.71
θ	0.54	0.035	<0.001	0.47	0.60
			E. Cumulative Prospect Theory		
α	0.21	0.021	<0.001	0.17	0.25
β	0.06	0.054	0.29	-0.05	0.16
λ	1.34	0.103	<0.001 (H_0: λ=1)	1.14	1.54
$\eta+$	1.04	0.052	<0.001	0.94	1.14
$\varphi+$	0.44	0.035	<0.001	0.37	0.51
$\eta-$	0.75	0.057	<0.001	0.63	0.86
$\varphi-$	0.97	0.039	<0.001	0.90	1.05

Figure 7: Prelec Probability Weighting and Implied Decision Weights for the Dual Theory Model

Based on equi-probable reference lotteries

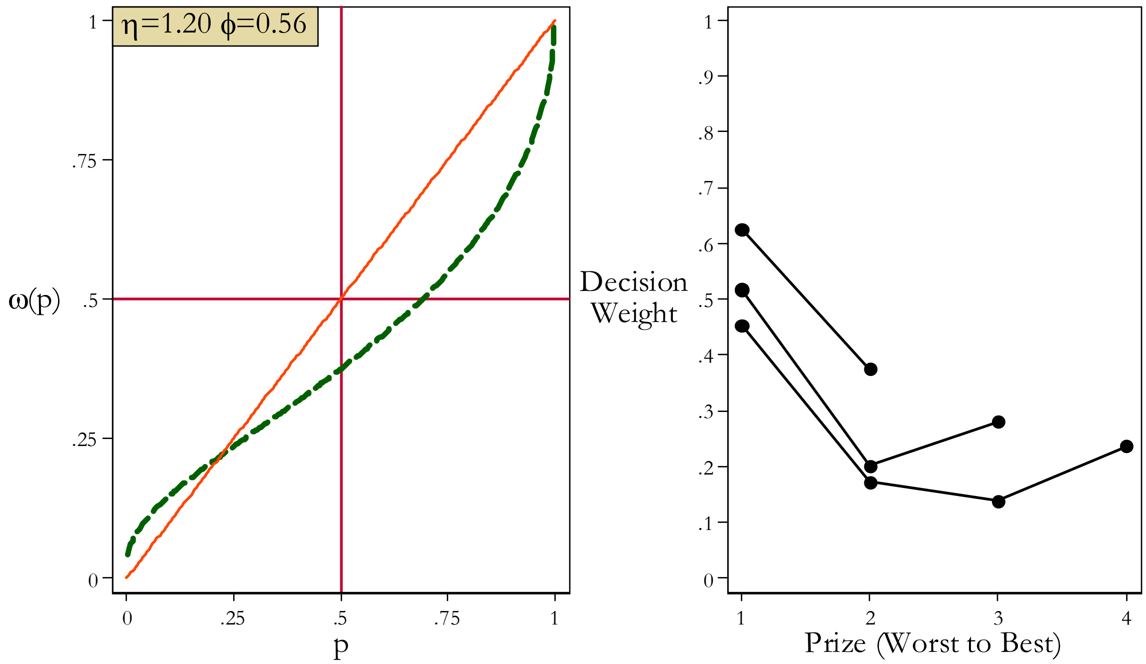


Figure 8: Prelec Probability Weighting and Implied Decision Weights for the Rank Dependent Model

Based on equi-probable reference lotteries

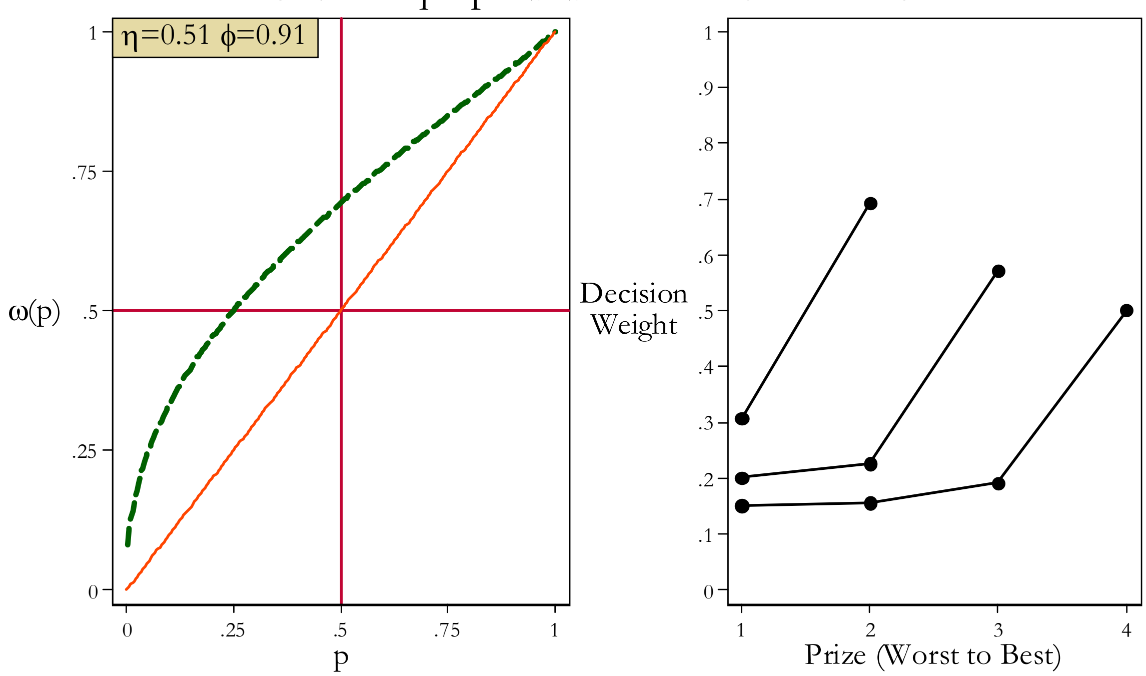


Figure 9: CPT Model for GSU Undergraduates with House Money to Cover Losses

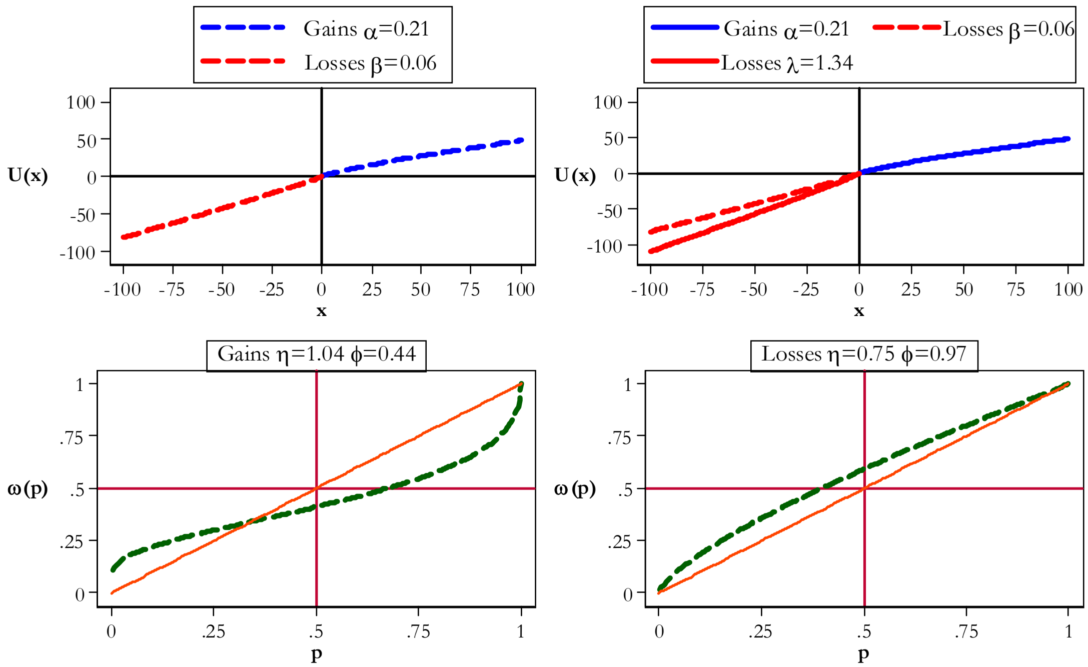


Figure 10: Probability Weighting and Decision Weights for the CPT Model with GSU Undergraduates

$\alpha = .21 \quad \beta = .057 \quad \lambda = 1.34$

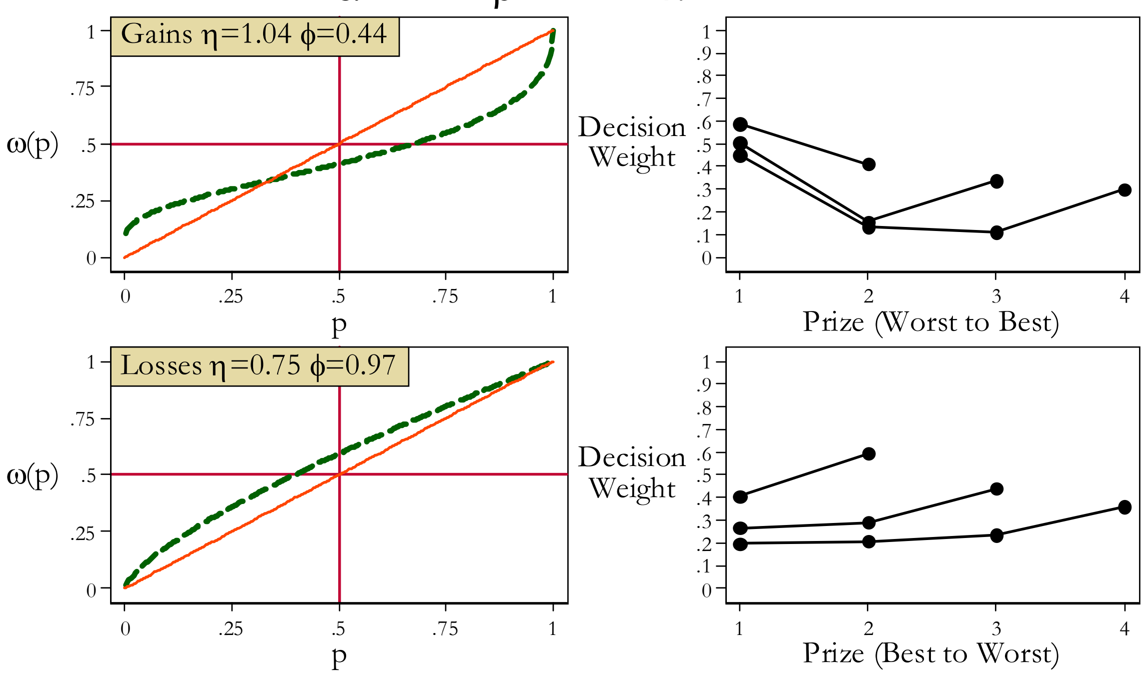


Figure 11: Classifying Subjects as EUT, DT, RDU or DA Using Hypothesis Test that ω(p)=p or θ=0 and a Significance Level of 5% to Reject EUT

N=177, one *p*-value per individual
Estimates for each individual of EUT, DT, RDU and DA models

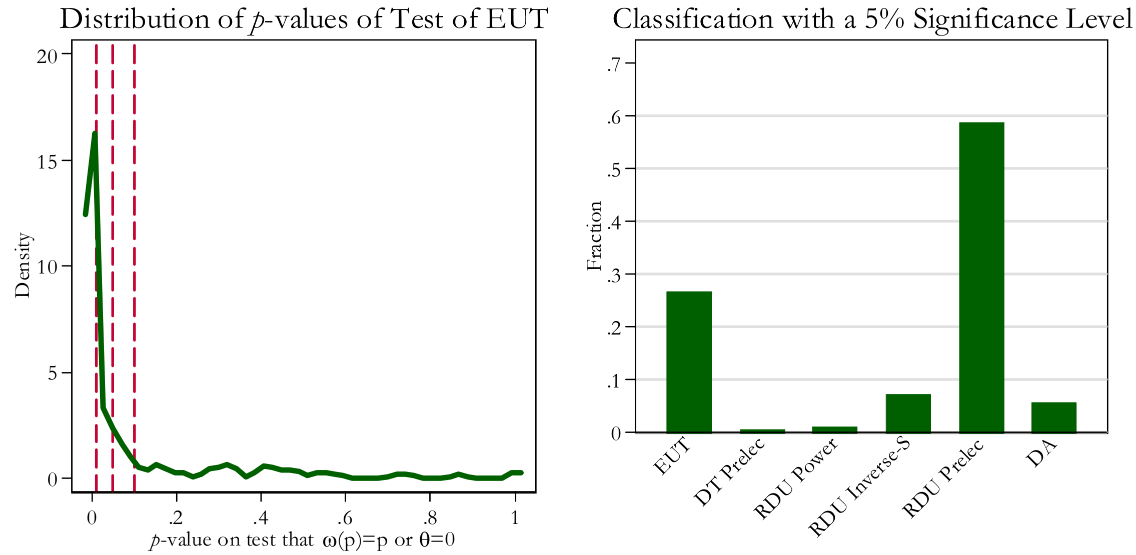


Figure 12: Distribution of Log Likelihood Ratios for Non-Nested Tests of **RDU Prelec** and **CPT Prelec** Models for Subject #2 Making 100 Choices

Kernel density estimate of data and normal density for comparison
Vuoung statistic favors RDU Prelec with *p*-value = 0.49 on null of non-discrimination
Clarke statistic favors RDU Prelec with *p*-value = 0.06 on null of non-discrimination
Chen-Shapiro test of assumption of a normal distribution has *p*-value < 0.01

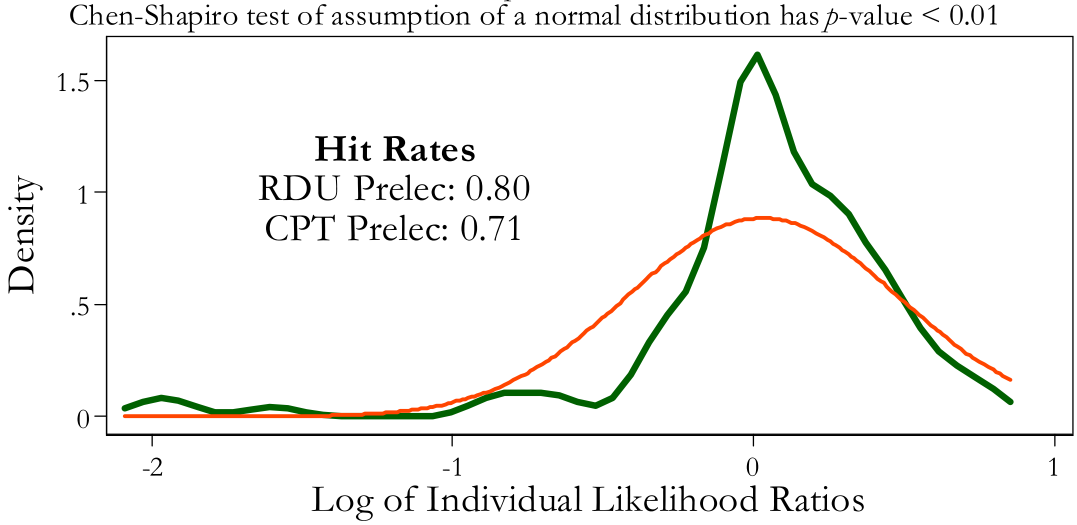


Figure 13: Distribution of Log Likelihood Ratios for Non-Nested Tests of **RDU Prelec** and **CPT Prelec** Models for Subject #7 Making 100 Choices

Kernel density estimate of data and normal density for comparison
Vuoung statistic favors CPT Prelec with *p*-value < 0.01 on null of non-discrimination
Clarke statistic favors CPT Prelec with *p*-value < 0.01 on null of non-discrimination
Chen-Shapiro test of assumption of a normal distribution has *p*-value = 0.04

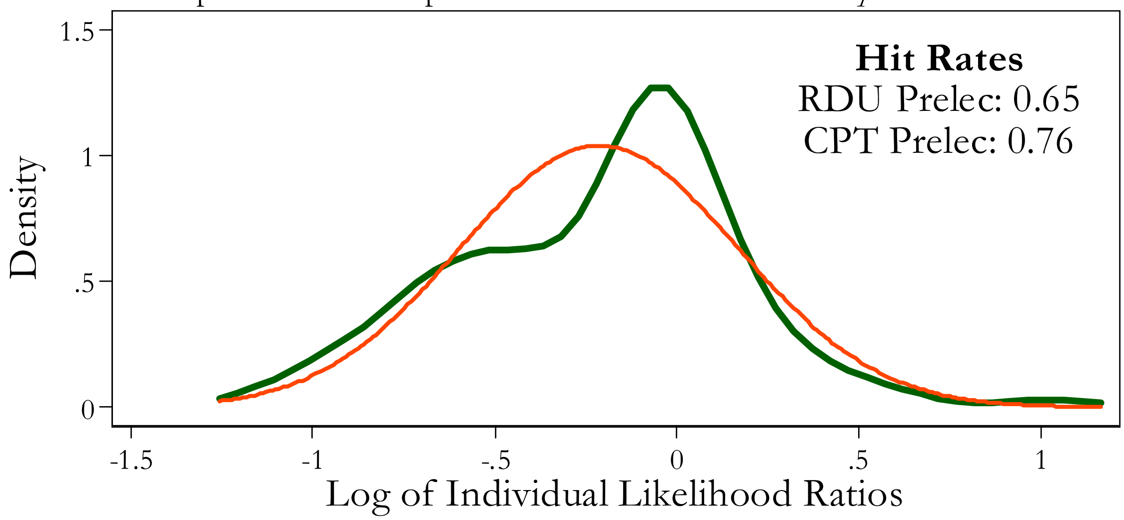


Figure 14: Distribution of Log Likelihood Ratios for Non-Nested Tests of **EUT** and **CPT Prelec** Models for Subject #8 Making 100 Choices

Kernel density estimate of data and normal density for comparison
Vuoung statistic favors EUT with *p*-value = 0.02 on null of non-discrimination
Clarke statistic favors EUT with *p*-value = 0.09 on null of non-discrimination
Chen-Shapiro test of assumption of a normal distribution has *p*-value < 0.01

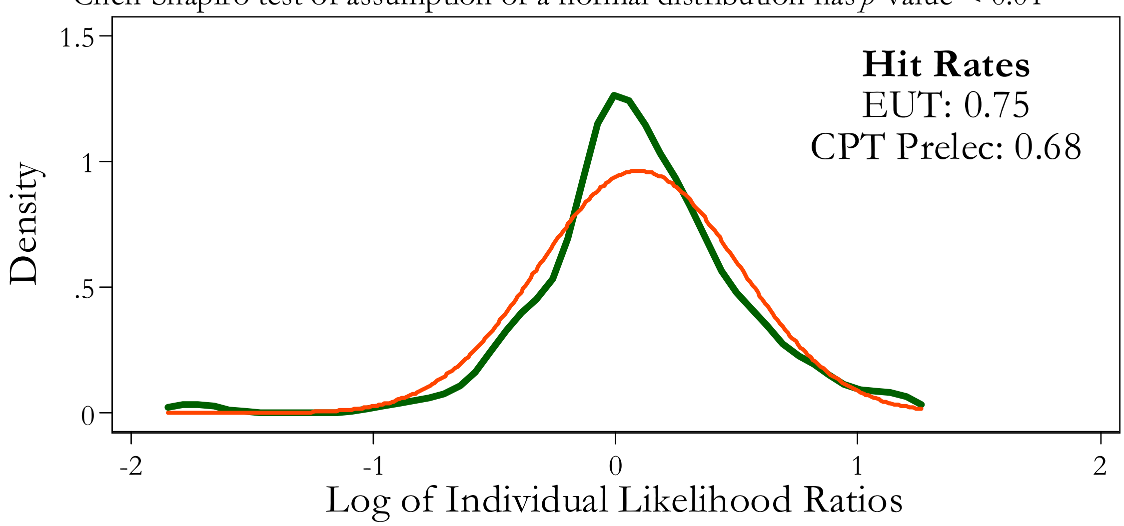


Figure 15: Distribution of Log Likelihood Ratios for
Non-Nested Tests of **EUT** and **CPT Prelec**
Models for Subject #3 Making 100 Choices

Kernel density estimate of data and normal density for comparison
Vuoung statistic favors CPT Prelec with p-value = 0.87 on null of non-discrimination
Clarke statistic favors CPT Prelec with p-value = 0.19 on null of non-discrimination
Chen-Shapiro test of assumption of a normal distribution has p-value < 0.01

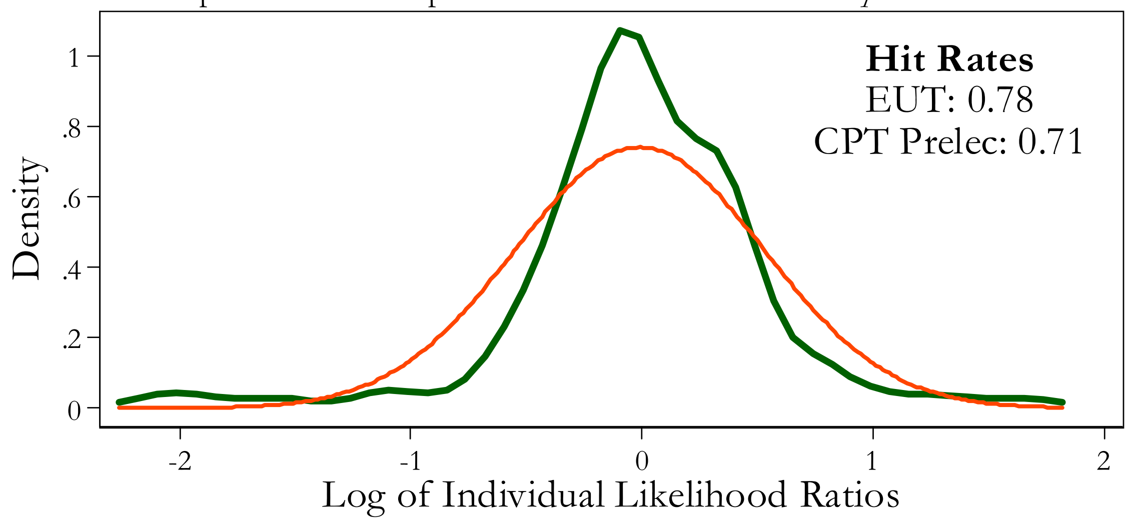


Figure 16: Share of Subjects Favoring One of the CPT Models

Direction of the Vuoung or Clarke Non-Nested hypothesis test statistic
No account for statistical signifance of that direction

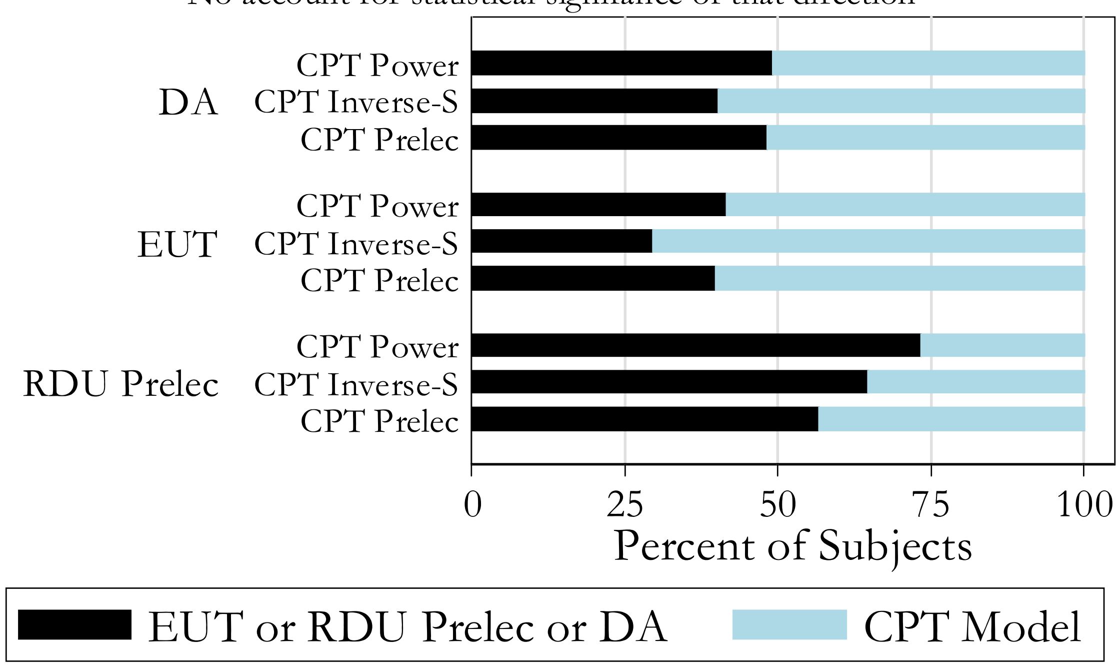


Figure 17: Average of Null Hypothesis *p*-values When CPT Model Favored

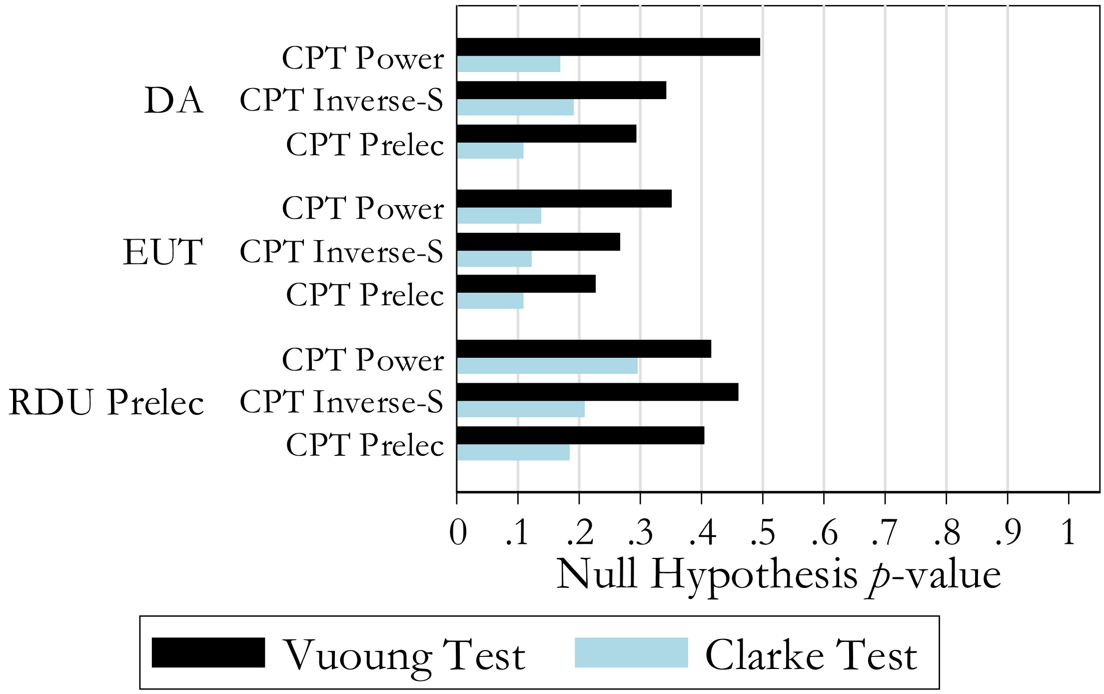


Figure 18: Average of Null Hypothesis *p*-values for Test of Normal Distribution of Vuoung Statistic When CPT Model Favored

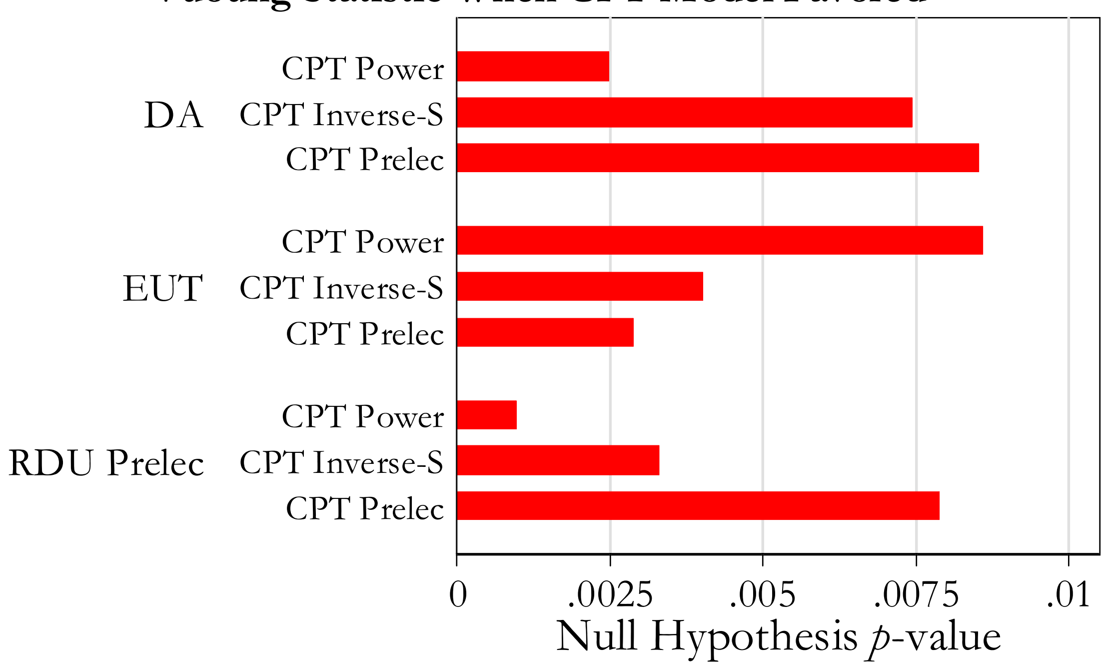


Figure 19: Distribution of *p*-values of Clarke Test of Hypothesis of Non-Discrimination for RDU Prelec and CPT Prelec Models

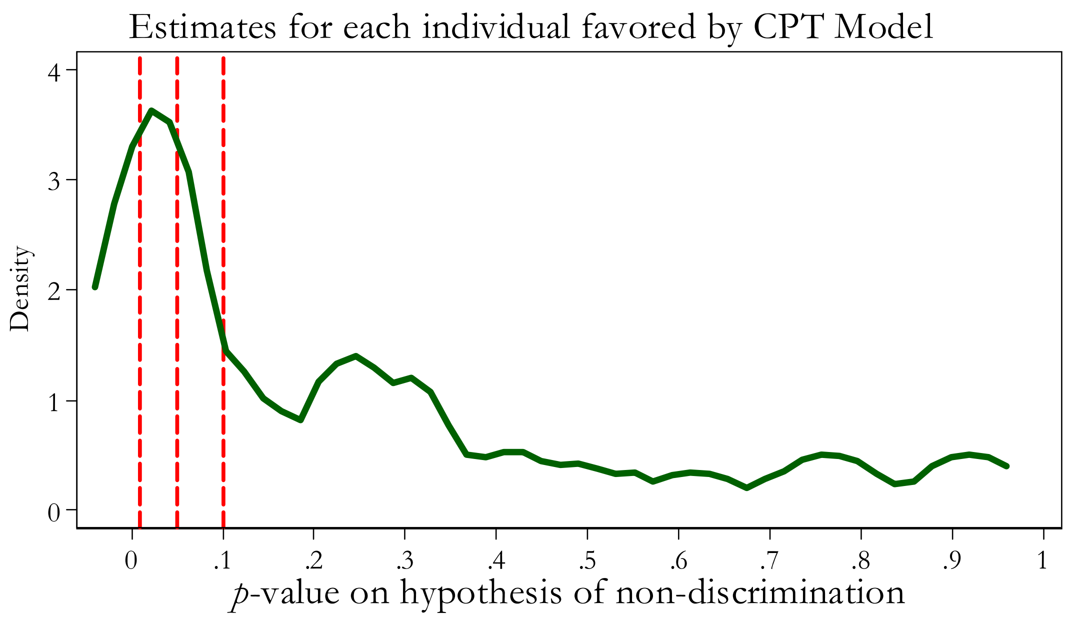


Figure 20: Classifying Subjects as EUT, DT, RDU, DA or CPT

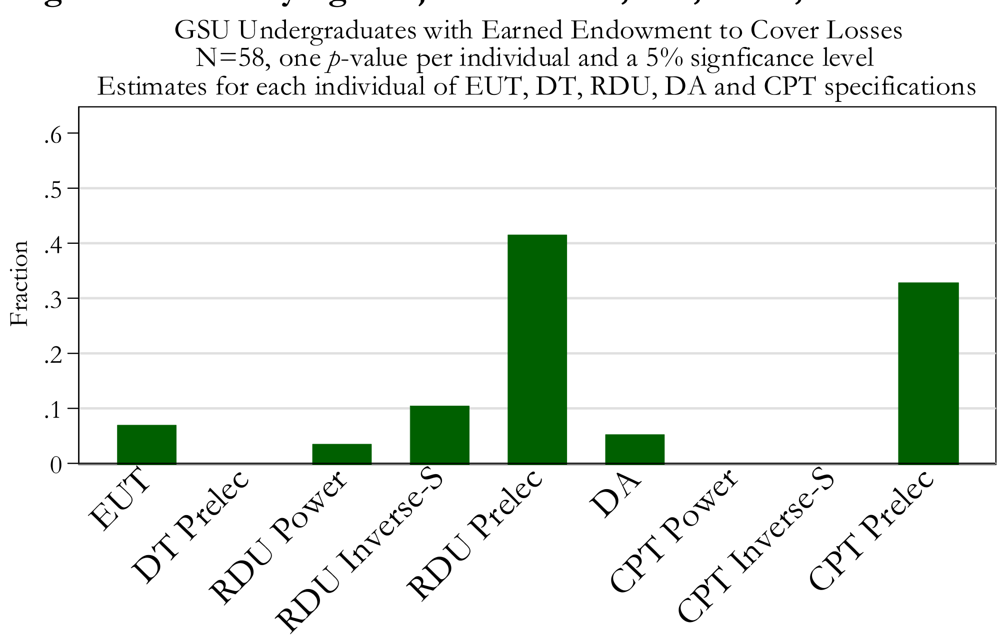


Table 2: The Existing Literature Claiming to Estimate CPT

Study	Rewards	Frames	Comments
Tversky and Kahneman [1992]	Non-salient	Gain, Loss	"Median" estimates reported.
Camerer and Ho [1994]	Real	Gain	
Wu and Gonzalez [1996]	Hypothetical	Gain	
Gonzalez and Wu [1999]	Real	Gain	
Fennema and van Assen [1998]	Hypothetical	Gain, Loss, Mixed	
Abdellaoui [2000]	Real Hypothetical	Gain Loss	
Schmidt and Traub [2002]	Hypothetical	Gain, Loss	
Harbaugh, Krause and Vesterlund [2002]	Real	Gain, Loss	Assumes no utility loss aversion. Claim to be unable to jointly estimate probability weighting and the value function.
Pennings and Smidts [2003]	Hypothetical	Gain, Loss[†]	
Etchart-Vincent [2004]	Hypothetical	Loss	
Mason, Shogren, Settle and List [2005]	Real	Loss	
Schunk and Betsch [2006]	Real Hypothetical	Gain Loss	
Stott [2006]	"Slightly Real"[¶]	Gain	Does not mention loss aversion.
Fehr-Duda, Gennaro, and Schubert [2006]	Real	Gain, Loss	Assumes no utility loss aversion.
Abdellaoui, Bleichrodt and l'Haridon [2008]	Real Hypothetical	Gain Loss, Mixed	
Rieskamp [2008]	"Slightly Real"[‡]	Gain, Loss, Mixed	Constrained to show loss aversion.
Booij and van de Kuilen [2009]	Hypothetical	Gain, Loss	

Booij, van Praag and van de Kuilen [2010]	Hypothetical	Gain, Loss, Mixed	
Bruhin, Fehr-Duda and Epper [2010]	Real	Gain, Loss	Assumes no utility loss aversion.
Pachur, Hanoch and Gummerum [2010]	Hypothetical	Gain, Loss Mixed	
von Gaudecker, van Soest and Wengström [2011]	Real	Gain, Mixed	Assumes no probability weighting.
Nilsson, Rieskamp and Wagenmakers [2011]	"Slightly Real" ‡	Gain, Loss, Mixed	
Glöckner and Pachur [2012]	Real	Gain, Loss, Mixed	"Median" estimates reported, apparently with no standard errors
Zeisberger, Vrecko and Langer [2012]	Real §	Gain, Loss, Mixed	Becker, DeGroot and Marschak [1964] method used to elicit certainty-equivalents.
Abdelloui, l'Haridon and Paraschiv [2013]	Real	Gain	
Scholten and Read [2014]	Non-Salient	Gain, Loss	Assumes no utility loss aversion.
Balcombe and Fraser [2016]		Gain	Does not mention loss aversion.
Bouchouicha and Vieder [2016]	Hypothetical	Gain, Loss	Assumes no utility loss aversion.

Notes: † Subject elicitations were all in the gain frame, but the authors' assumed (p. 1254) some positive reference point in their analysis and treated gains below that as "losses" for the purposes of analysis.

‡ Subjects made binary choices over lotteries with outcomes between +€100 and -€100, one of 180 choices was selected for payment and realization, and then 5% of the outcome added or subtracted from an endowment of €15.

¶ Lottery prizes up to £40,000 were included in binary lottery choices. Each subject was given a fixed £3, and one of the 90 choices selected, re-scaled so that the maximum prize would be £5, and then played out.

§ One subject in ten was selected for payment, but the losses in that case were substantial (up to €60) out of an endowment that had been earned in a previous task in that session.

Figure 21: Probability Weighting and Decision Weights from Mode of Bayesian Posterior Distributions Estimated by Nilsson, Rieskamp and Wagenmakers [2011]

Based on equi-probable reference lotteries

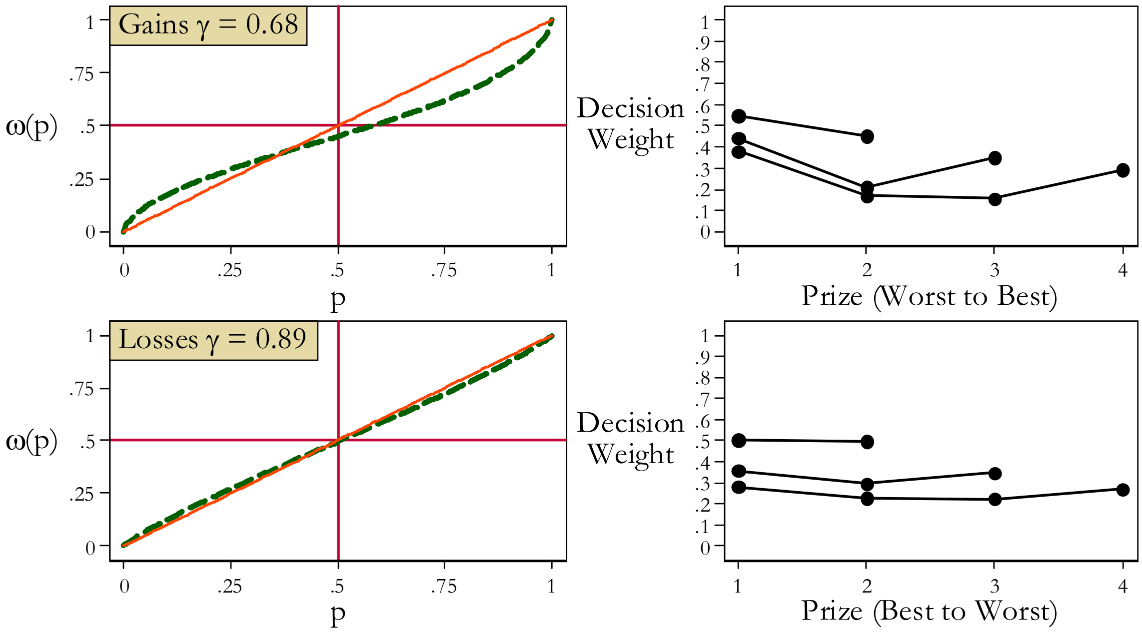
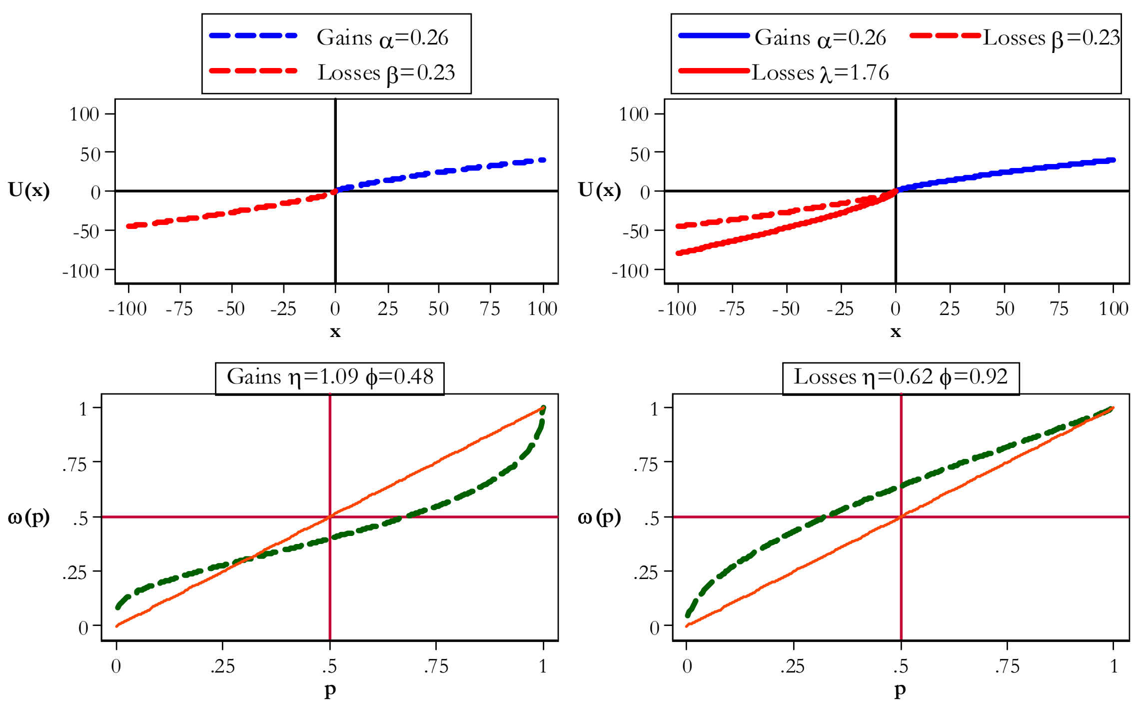


Figure 22: CPT Model for GSU Undergraduates with Earned Endowments to Cover Losses



References

Abdellaoui, Mohammed, "Parameter-Free Elicitation of Utilities and Probability Weighting Functions," *Management Science*, 46, 2000, 1497-1512.

Abdellaoui, Mohammed and Bleichrodt, Han, "Eliciting Gul's Theory of Disappointment Aversion by the Tradeoff Method," *Journal of Economic Psychology*, 28, 2007, 631-645.

Abdellaoui, Mohammed; Bleichrodt, Han, and Paraschiv, Corina, "Measuring Loss Aversion under Prospect Theory: A Parameter-Free Approach," *Management Science*, 53(10), October 2007, 1659-1674.

Abdellaoui, Mohammed; Bleichrodt, Han, and l'Haridon, Olivier, "A Tractable Method to Measure Utility and Loss Aversion under Prospect Theory," *Journal of Risk and Uncertainty*, 36, 2008, 245-266.

Abdellaoui, Mohammed; l'Haridon, Olivier, and Paraschiv, Corina, "Individual vs. Couple Behavior: An Experimental Investigation of Risk Preferences," *Theory and Decision*, 75(2), 2013, 175-191.

Andersen, Steffen; Harrison, Glenn W.; Hole, Arne Rise; Lau, Morten I., and Rutström, E. Elisabet, "Non-Linear Mixed Logit," *Theory and Decision*, 73, 2012, 77-96.

Andersen, Steffen; Harrison, Glenn W.; Lau, Morten I., and Rutström, E. Elisabet, "Elicitation Using Multiple Price Lists," *Experimental Economics*, 9(4), December 2006, 383-405.

Balcombe, Kelvin, and Fraser, Iain, "Parametric Preference Functionals Under Risk in the Gain Domain: A Bayesian Analysis," *Journal of Risk & Uncertainty*, 50, 2015, 161-187.

Becker, Gordon M.; DeGroot, Morris H., and Marschak, Jacob., "Measuring Utility By A Single-Response Sequential Method," *Behavioral Science*, 9, July 1964, 226-232.

Billor, Nedret; Ali S. Hadi, and Paul F. Velleman, "BACON: blocked adaptive computationally efficient outlier nominators," *Computational Statistics & Data Analysis*, 34, 2000, 279-298.

Bleichrodt, Han; Pinto, J.L, and Wakker, Peter P., "Using Descriptive Findings of Prospect Theory to Improve the Prescriptive Use of Expected Utility," *Management Science*, 47, 2001, 1498-1514.

Booij, Adam S., and van de Kuilen, Gijs, "A Parameter-Free Analysis of the Utility of Money for the General Population Under Prospect Theory," *Journal of Economic Psychology*, 30, 2009, 651-666.

Booij, Adam S.; van Praag, Bernard M.S., and van de Kuilen, Gijs, "A Parametric Analysis of Prospect Theory's Functionals for the General Population," *Theory and Decision*, 68, 2010, 115-148.

Bosh-Domènech, Antoni, and Silvestre, Joaquim, "Averting Risk in the Face of Large Losses: Bernoulli vs. Tversky and Kahnman," *Economics Letters*, 107, 2010, 180-182.

Bouchouicha, Ranoua and Vieider, Ferdinand M., "Accommodating Stake Effects under Prospect Theory," *Discussion Paper 2016-120*, Department of Economics, University of Reading, 2016.

Brooks, Peter; Peters, Simon, and Zank, Horst, "Risk Behavior for Gain, Loss, and Mixed Prospects," *Theory and Decision*, 77, 2014, 153-182.

Brooks, Peter, and Zank, Horst, "Loss Averse Behavior," *Journal of Risk & Uncertainty*, 31(3), 2005, 301-325.

Bruhin, Adrian; Fehr-Duda, and Epper, Thomas, "Risk and Rationality: Uncovering Heterogeneity in Probability Distortion," *Econometrica*, 78(4), July 2010, 1375-1412.

Brzezinski, Michal, "The Chen-Shapiro Test for Normality," *The Stata Journal*, 2012, 12(3), 368-374.

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

Camerer, Colin F., "Recent Tests of Generalizations of Expected Utility Theory," in W. Edwards (ed.), *Utility Theories: Measurements and Applications* (Boston: Kluwer, 1992).

Camerer, Colin F., "Three Cheers – Psychological, Theoretical, Empirical – for Loss Aversion," *Journal of Marketing Research*, XLII, May 2005, 129-133.

Camerer, Colin, and Ho, Teck-Hua, "Violations of the Betweenness Axiom and Nonlinearity in Probability," *Journal of Risk & Uncertainty*, 8, 1994, 167-196.

Cárdenas, Juan Camilo; De Roux, Nicolas; Jaramillo, Christian R., and Martinez, Luis Roberto, "Is it My Money or Not? An Experiment on Risk Aversion and the House-money Effect," *Experimental Economics*, 17, 2014, 47-60.

Chateauneuf, Alain, and Wakker, Peter, "An Axiomatization of Cumulative Prospect Theory for Decisions Under Risk," *Journal of Risk and Uncertainty*, 18(2), 1999, 137-145.

Chen, Ling, and Shapiro, Samuel S., "An Alternative Test for Normality Based on Normalized Spacings," *Journal of Statistical Computation and Simulation*, 1995, 53, 269-288.

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

Clarke, Kevin A., "Nonparametric Model Discrimination in International Relations," *Journal of Conflict Resolution*, 47(1), February 2003, 72-93.

Clarke, Kevin A., "A Simple Distribution-Free Test for Non-Nested Model Selection," *Political Analysis*, 15(3), 2007, 347-363.

Clarke, Kevin A., and Signorino, Curtis S., "Discriminating Methods: Tests for Non-nested Discrete Choice Models," *Political Studies*, 2010, 58, 368-388.

Conlisk, John, "Three Variants on the Allais Example," *American Economic Review*, 79(3), June 1989, 392-407.

Cox, David R., "Tests of Separate Families of Hypotheses," in E.G. Charatsis (ed.), *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* (Berkeley: University of California Press, Volume 1, 1961, 105-123).

Cox, David R., "Further Results on Tests of Separate Families of Hypotheses," *Journal of the Royal Statistical Society, Series B*, 24, 1962, 406-424.

Cox, James C., and Sadiraj, Vjollca, "Small- and Large-Stakes Risk Aversion: Implications of Concavity Calibration for Decision Theory," *Games and Economic Behavior*, 56, 2006, 45-60.

Cox, James C., and Sadiraj, Vjollca, "Risky Decisions in the Large and in the Small: Theory and Experiment," in J. Cox and G.W. Harrison (eds.), *Risk Aversion in Experiments* (Bingley, UK: Emerald, Research in Experimental Economics, Volume 12, 2008).

Cox, James C.; Sadiraj, Vjollca, and Schmidt, Ulrich, "Paradoxes and Mechanisms for Choice under Risk," *Experimental Economics*, 18(2), June 2015, 215-250.

Cubitt, Robin P.; Starmer, Chris, and Sugden, Robert, "On the Validity of the Random Lottery Incentive System," *Experimental Economics*, 1(2), 1998, 115-131.

Eeckhoudt, Louis, and Schlesinger, Harris, "Putting Risk in Its Proper Place," *American Economic Review*, 96, 2006, 280-289.

El-Gamal, Mahmoud A., and Grether, David M., "Are People Bayesian? Uncovering Behavioral Strategies," *Journal of the American Statistical Association*, 90, 432, December 1995, 1137-1145.

Etchart-Vincent, Nathalie, "Is Probability Weighting Sensitive to the Magnitude of Consequences? An Experimental Investigation on Losses," *Journal of Risk & Uncertainty*, 28, 2004, 217-235.

Fehr-Duda, Helga; Gennaro, Manuelle, and Schubert, Renate, "Gender, Financial Risk, and Probability Weights," *Theory and Decision*, 60, 2006, 283-313.

Fennema, Hein, and van Assen, Marcel, "Measuring the Utility of Losses by Means of the Trade-off Method," *Journal of Risk & Uncertainty*, 17, 1998, 277-295.

Fishburn, Peter C., and Kochenberger, Gary A,. "Two-Piece von Neumann-Morgenstern Utility Functions," *Decision Sciences*, 10, 1979, 503-518.

Glöckner, Andreas, and Betsch, Tilmann, "Do People Male Decisions Under Risk Based on Ignorance? An Empirical Test of the Priority Heuristic Against Cumulative Prospect Theory," *Organizational Behavior and Human Decision Processes*, 107, 2008, 75-95.

Glöckner, Andreas, and Pachur, Thorsten, "Cognitive Models of Risky Choice: Parameter Stability and Predictive Accuracy of Prospect Theory," *Cognition*, 123(1), 2012, 21-32.

Gonzalez, Richard, and Wu, George, "On the Shape of the Probability Weighting Function," *Cognitive Psychology*, 38, 1999, 129-166.

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

Hansson, Bengt, "Risk Aversion as a Problem of Conjoint Measurement," in P. Gardenfors and N-E. Sahlin (eds.), *Decisions, Probability, and Utility* (New York: Cambridge University Press, 1988).

Harbaugh, William T.; Krause, Kate, and Vesterlund, Lise, "Risk Attitudes of Children and Adults: Choices over Small and Large Probability Gains and Losses," *Experimental Economics*, 5, 2002, 53-84.

Harless, David W., and Camerer, Colin F., "The Predictive Utility of Generalized Expected Utility Theories," *Econometrica*, 62(6), November 1994, 1251-1289.

Harrison, Glenn W.; Johnson, Eric; McInnes, Melayne M., and Rutström, E. Elisabet, "Measurement With Experimental Controls," in M. Boumans (ed.), *Measurement in Economics: A Handbook* (San Diego, CA: Elsevier, 2007).

Harrison, Glenn W.; Lau, Morten; Ross, Don; and Swarthout, J. Todd, "Small-Stakes Risk Aversion in the Laboratory: A Reconsideration," *CEAR Working Paper 2016-04*, Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University, 2016.

Harrison, Glenn W., and List, John A., "Field Experiments," *Journal of Economic Literature*, 42(4), December 2004, 1013-1059.

Harrison, Glenn W., and Rutström, E. Elisabet, "Risk Aversion in the Laboratory," in J.C. Cox and G.W. Harrison (eds.), *Risk Aversion in Experiments* (Bingley, UK: Emerald, Research in Experimental Economics, Volume 12, 2008).

Harrison, Glenn W., and Rutström, E. Elisabet, "Expected Utility *And* Prospect Theory: One Wedding and A Decent Funeral," *Experimental Economics*, 12(2), June 2009, 133-158.

Harrison, Glenn W., and Swarthout, J. Todd, "Experimental Payment Protocols and the Bipolar Behaviorist," *Theory and Decision*, 77(3), 2014, 423-438.

Hey, John D., and Orme, Chris, "Investigating Generalizations of Expected Utility Theory Using Experimental Data," *Econometrica*, 62(6), November 1994, 1291-1326.

Holt, Charles A., "Preference Reversals and the Independence Axiom," *American Economic Review*, 76, June 1986, 508-514.

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

Karni, Edi, and Safra, Zvi, "Preference Reversals and the Observability of Preferences by Experimental Methods," *Econometrica*, 55, 1987, 675-685.

Köbberling, Veronika, and Wakker, Peter P., "An Index of Loss Aversion," *Journal of Economic Theory*, 122, 2005, 119-131.

Kőszegi, Botond, and Rabin, Matthew, "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics*, 121(4), November 2006, 1133-1165.

Kőszegi, Botond, and Rabin, Matthew, "Reference-Dependent Risk Attitudes," *American Economic Review*, 97(4), September 2007, 1047-1073.

Laury, Susan K.; McInnes, Melayne M., and Swarthout, J. Todd, "Insurance Decisions for Low-Probability Losses," *Journal of Risk and Uncertainty*, 39, 2009, 17-44.

Loomes, Graham, and Sugden, Robert, "Testing Different Stochastic Specifications of Risky Choice," *Economica*, 65, 1998, 581-598.

Luce, R. Duncan, and Fishburn, Peter C., "Rank and Sign-Dependent Linear Utility Models for Finite First-Order Gambles," *Journal of Risk & Uncertainty*, 4, 1991, 29-59.

Mason, Charles F.; Shogren, Jason F.; Settle, Chad, and List, John A., "Investigating Risky Choices Over Losses Using Experimental Data," *Journal of Risk and Uncertainty*, 31(2), 187-215, 2005.

Neilson, William, and Stowe, Jill, "A Further Examination of Cumulative Prospect Theory Paramaterizations," *Journal of Risk and Uncertainty*, 24(1), 2002, 31-46.

Nilsson, Håkan; Rieskamp, Jörg, and Wagenmakers, Eric-Jan, "Hierarchical Bayesian Parameter Estimation for Cumulative Prospect Theory," *Journal of Mathematical Psychology*, 55, 2011, 84-93.

Pachur, Thorsten; Hanoch, Yaniv, and Gummerum, Michaela, "Prospects Behind Bars: Analyzing Decisions Under Risk in a Prison Population," *Psychonomic Bulletin and Review*, 17, 2010, 630-636.

Pennings, Joost M.E, and Smidts, Ale, "The Shape of Utility Functions and Organizational Behavior," *Management Science*, 24, 2003, 1251-1263.

Prelec, Drazen, "The Probability Weighting Function," *Econometrica*, 66, 1998, 497-527.

Quiggin, John, "A Theory of Anticipated Utility," *Journal of Economic Behavior & Organization*, 3(4), 1982, 323-343.

Rabin, Matthew, "Risk Aversion and Expected Utility Theory: A Calibration Theorem," *Econometrica*, 68, 2000, 1281-1292.

Ray, Debajyoti; Shum, Matthew, and Camerer, Colin F., "Loss Aversion in Post-Sale Purchases of Consumer Products and their Substitutes," *American Economic Review (Papers & Proceedings)*, 105(5), May 2015, 376-380.

Rieskamp, Jörg, "The Probabilistic Nature of Preferential Choice," *Journal of Experimental Psychology: Learning, Memory and Cognition*, 34(6), 2008, 1446-1465.

Saha, Atanu, "Expo-Power Utility: A Flexible Form for Absolute and Relative Risk Aversion," *American Journal of Agricultural Economics*, 75(4), November 1993, 905-913.

Scholten, Marc, and Read, Daniel, "Prospect Theory and the 'Forgotten' Fourfold Pattern of Risk Preferences," *Journal of Risk & Uncertainty*, 48(1), 2014, 67-83.

Schmidt, Ulrich; Starmer, Chris, and Sugden, Robert, "Third-Generation Prospect Theory," *Journal of Risk & Uncertainty*, 36(3), June 2008, 203-223.

Schmidt, Ulrich, and Traub, Stefan, "An Experimental Test of Loss Aversion," *Journal of Risk & Uncertainty*, 25, 2002, 233-249.

Schmidt, Ulrich, and Zank, Horst, "Risk Aversion in Cumulative Prospect Theory," *Management Science*, 54, 2008, 208–216.

Schmidt, Ulrich, and Zank, Horst, "A Simple Model of Cumulative Prospect Theory," *Journal of Mathematical Economics*, 45(3-4), March 2009, 308–319.

Schunk, Daniel, and Betsch, Cornelia, "Explaining Heterogeneity in Utility Functions by Individual Differences in Decision Modes," *Journal of Economic Psychology*, 27, 2006, 386-401.

Stahl, Dale O., and Wilson, Paul W., "On Players' Models of Other Players: Theory and Experimental Evidence," *Games and Economic Behavior*, 10, 1995, 218-254.

Starmer, Chris, and Sugden, Robert, "Violations of the Independence Axiom in Common Ratio Problems: An Experimental Test of Some Competing Hypotheses," *Annals of Operational Research*, 19, 1989, 79-102.

Starmer, Chris, "Developments in Non-Expected Utility Theory: The Hunt for a Descriptive Theory of Choice Under Risk," *Journal of Economic Literature*, 38, June 2000, 332-382.

Starmer, Chris, and Sugden, Robert, "Does the Random-Lottery Incentive System Elicit True Preferences? An Experimental Investigation," *American Economic Review*, 81, 1991, 971-978.

Stott, Henry P., "Cumulative Prospect Theory's Functional Menagerie," *Journal of Risk and Uncertainty*, 32, 2006, 101-130.

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

von Gaudecker, Hans-Martin; van Soest, Arthur, and Wengström, Erik, "Heterogeneity in Risky Choice Behavior in a Broad Population," *American Economic Review*, 101, April 2011, 664-694.

Vuong, Quang H., "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses," *Econometrica*, 57(2), March 1989, 307-333.

Wakker, Peter P., *Prospect Theory for Risk and Ambiguity* (New York: Cambridge University Press, 2010).

Wakker, Peter, and Tversky, Amos, "An Axiomatization of Cumulative Prospect Theory," *Journal of Risk and Uncertainty*, 7(7), 1993, 147-176.

Wilcox, Nathaniel T., "Stochastic Models for Binary Discrete Choice Under Risk: A Critical Primer and Econometric Comparison," in J. Cox and G.W. Harrison (eds.), *Risk Aversion in Experiments* (Bingley, UK: Emerald, Research in Experimental Economics, Volume 12, 2008).

Wilcox, Nathaniel T., "'Stochastically More Risk Averse:' A Contextual Theory of Stochastic Discrete Choice Under Risk," *Journal of Econometrics*, 162(1), May 2011, 89-104.

Wilcox, Nathaniel T., "Is the Premise of Risk Calibration Theorems Plausible?" *Presentation*, CEAR Workshop, Durham University, September 17, 2013.

Wooldridge, Jeffrey, "Cluster-Sample Methods in Applied Econometrics," *American Economic Review (Papers & Proceedings)*, 93, May 2003, 133-138.

Wu, George, and Gonzalez, Richard, "Curvature of the Probability Weighting Function," *Management Science*, 42, 1996, 1676-1690.

Yaari, Menahem E., "The Dual Theory of Choice under Risk," *Econometrica*, 55(1), 1987, 95-115.

Zeisberger, Stefan; Vrecko, Dennis, and Langer, Thomas, "Measuring the Time Stability of Prospect Theory Preferences," *Theory and Decision*, 72, 2012, 359-386.

Appendix A: Parameters of Experiments

Table A1: Battery of 100 Lottery Tasks in Choices Made by Undergraduate Subjects

Task	EV left	EV right	EV ratio	Left $ 1	Left p 1	Left $ 2	Left p 2	Left $ 3	Left p 3	Right $ 1	Right p 1	Right $ 2	Right p 2	Right $ 3	Right p 3	Notes
1	$5.00	$6.95	-28%	$0	0	$5	1	$0	0	$0	0.01	$5	0.89	$25	0.1	Allais - lower stakes
2	$0.55	$2.50	-78%	$0	0.89	$5	0.11	$0	0	$0	0.9	$5	0	$25	0.1	Allais - lower stakes
3	$15.00	$20.85	-28%	$0	0	$15	1	$0	0	$0	0.01	$15	0.89	$75	0.1	Allais - higher stakes
4	$1.65	$7.50	-78%	$0	0.89	$15	0.11	$0	0	$0	0.9	$15	0	$75	0.1	Allais - higher stakes
5	$59.50	$61.25	-3%	$0	0.15	$35	0	$70	0.85	$0	0	$35	0.25	$70	0.75	LS1: Loomes and Sugden
6	$49.00	$50.75	-3%	$0	0.3	$35	0	$70	0.7	$0	0.15	$35	0.25	$70	0.6	LS2: Loomes and Sugden
7	$49.00	$52.50	-7%	$0	0.3	$35	0	$70	0.7	$0	0	$35	0.5	$70	0.5	LS3: Loomes and Sugden
8	$50.75	$52.50	-3%	$0	0.15	$35	0.25	$70	0.6	$0	0	$35	0.5	$70	0.5	LS4: Loomes and Sugden
9	$33.25	$35.00	-5%	$0	0.15	$35	0.75	$70	0.1	$0	0	$35	1	$70	0	LS5: Loomes and Sugden
10	$28.00	$35.00	-20%	$0	0.6	$35	0	$70	0.4	$0	0	$35	1	$70	0	LS6: Loomes and Sugden
11	$28.00	$33.25	-16%	$0	0.6	$35	0	$70	0.4	$0	0.15	$35	0.75	$70	0.1	LS7: Loomes and Sugden
12	$7.00	$8.75	-20%	$0	0.9	$35	0	$70	0.1	$0	0.75	$35	0.25	$70	0	LS8: Loomes and Sugden
13	$63.00	$63.00	0%	$0	0.1	$35	0	$70	0.9	$0	0	$35	0.2	$70	0.8	LS9: Loomes and Sugden
14	$35.00	$35.00	0%	$0	0.5	$35	0	$70	0.5	$0	0.1	$35	0.8	$70	0.1	LS10: Loomes and Sugden
15	$35.00	$35.00	0%	$0	0.5	$35	0	$70	0.5	$0	0	$35	1	$70	0	LS11: Loomes and Sugden
16	$35.00	$35.00	0%	$0	0.1	$35	0.8	$70	0.1	$0	0	$35	1	$70	0	LS12: Loomes and Sugden
17	$21.00	$21.00	0%	$0	0.7	$35	0	$70	0.3	$0	0.5	$35	0.4	$70	0.1	LS13: Loomes and Sugden
18	$21.00	$21.00	0%	$0	0.7	$35	0	$70	0.3	$0	0.4	$35	0.6	$70	0	LS14: Loomes and Sugden
19	$21.00	$21.00	0%	$0	0.5	$35	0.4	$70	0.1	$0	0.4	$35	0.6	$70	0	LS15: Loomes and Sugden
20	$7.00	$7.00	0%	$0	0.9	$35	0	$70	0.1	$0	0.8	$35	0.2	$70	0	LS16: Loomes and Sugden
21	$63.00	$61.25	3%	$0	0.1	$35	0	$70	0.9	$0	0	$35	0.25	$70	0.75	LS17: Loomes and Sugden
22	$42.00	$36.75	14%	$0	0.4	$35	0	$70	0.6	$0	0.1	$35	0.75	$70	0.15	LS18: Loomes and Sugden
23	$42.00	$35.00	20%	$0	0.4	$35	0	$70	0.6	$0	0	$35	1	$70	0	LS19: Loomes and Sugden
24	$36.75	$35.00	5%	$0	0.1	$35	0.75	$70	0.15	$0	0	$35	1	$70	0	LS20: Loomes and Sugden
25	$21.00	$19.25	9%	$0	0.7	$35	0	$70	0.3	$0	0.6	$35	0.25	$70	0.15	LS21: Loomes and Sugden
26	$21.00	$17.50	20%	$0	0.7	$35	0	$70	0.3	$0	0.5	$35	0.5	$70	0	LS22: Loomes and Sugden
27	$19.25	$17.50	10%	$0	0.6	$35	0.25	$70	0.15	$0	0.5	$35	0.5	$70	0	LS23: Loomes and Sugden
28	$10.50	$8.75	20%	$0	0.85	$35	0	$70	0.15	$0	0.75	$35	0.25	$70	0	LS24: Loomes and Sugden
29	$63.00	$59.50	6%	$0	0.1	$35	0	$70	0.9	$0	0	$35	0.3	$70	0.7	LS25: Loomes and Sugden
30	$42.00	$35.00	20%	$0	0.4	$35	0	$70	0.6	$0	0.2	$35	0.6	$70	0.2	LS26: Loomes and Sugden
31	$42.00	$31.50	33%	$0	0.4	$35	0	$70	0.6	$0	0.1	$35	0.9	$70	0	LS27: Loomes and Sugden
32	$35.00	$31.50	11%	$0	0.2	$35	0.6	$70	0.2	$0	0.1	$35	0.9	$70	0	LS28: Loomes and Sugden
33	$28.00	$24.50	14%	$0	0.6	$35	0	$70	0.4	$0	0.5	$35	0.3	$70	0.2	LS29: Loomes and Sugden
34	$28.00	$21.00	33%	$0	0.6	$35	0	$70	0.4	$0	0.4	$35	0.6	$70	0	LS30: Loomes and Sugden
35	$24.50	$21.00	17%	$0	0.5	$35	0.3	$70	0.2	$0	0.4	$35	0.6	$70	0	LS31: Loomes and Sugden
36	$14.00	$10.50	33%	$0	0.8	$35	0	$70	0.2	$0	0.7	$35	0.3	$70	0	LS32: Loomes and Sugden
37	$63.00	$56.00	13%	$0	0.1	$35	0	$70	0.9	$0	0	$35	0.4	$70	0.6	LS33: Loomes and Sugden
38	$52.50	$42.00	25%	$0	0.25	$35	0	$70	0.75	$0	0.1	$35	0.6	$70	0.3	LS34: Loomes and Sugden
39	$52.50	$35.00	50%	$0	0.25	$35	0	$70	0.75	$0	0	$35	1	$70	0	LS35: Loomes and Sugden
40	$42.00	$35.00	20%	$0	0.1	$35	0.6	$70	0.3	$0	0	$35	1	$70	0	LS36: Loomes and Sugden

41	$28.00	$21.00	33%	$0	0.5	$35	0.2	$70	0.3	$0	0.4	$35	0.6	$70	0	LS37: Loomes and Sugden
42	$31.50	$21.00	50%	$0	0.55	$35	0	$70	0.45	$0	0.4	$35	0.6	$70	0	LS38: Loomes and Sugden
43	$31.50	$28.00	13%	$0	0.55	$35	0	$70	0.45	$0	0.5	$35	0.2	$70	0.3	LS39: Loomes and Sugden
44	$21.00	$14.00	50%	$0	0.7	$35	0	$70	0.3	$0	0.6	$35	0.4	$70	0	LS40: Loomes and Sugden
45	($63.00)	($63.00)	0%	$0	0.1	($35)	0	($70)	0.9	$0	0	($35)	0.2	($70)	0.8	LS9: Loomes and Sugden
46	($35.00)	($35.00)	0%	$0	0.5	($35)	0	($70)	0.5	$0	0.1	($35)	0.8	($70)	0.1	LS10: Loomes and Sugden
47	($35.00)	($35.00)	0%	$0	0.5	($35)	0	($70)	0.5	$0	0	($35)	1	($70)	0	LS11: Loomes and Sugden
48	($35.00)	($35.00)	0%	$0	0.1	($35)	0.8	($70)	0.1	$0	0	($35)	1	($70)	0	LS12: Loomes and Sugden
49	($21.00)	($21.00)	0%	$0	0.7	($35)	0	($70)	0.3	$0	0.5	($35)	0.4	($70)	0.1	LS13: Loomes and Sugden
50	($21.00)	($21.00)	0%	$0	0.7	($35)	0	($70)	0.3	$0	0.4	($35)	0.6	($70)	0	LS14: Loomes and Sugden
51	($21.00)	($21.00)	0%	$0	0.5	($35)	0.4	($70)	0.1	$0	0.4	($35)	0.6	($70)	0	LS15: Loomes and Sugden
52	($7.00)	($7.00)	0%	$0	0.9	($35)	0	($70)	0.1	$0	0.8	($35)	0.2	($70)	0	LS16: Loomes and Sugden
53	($63.00)	($56.00)	13%	$0	0.1	($35)	0	($70)	0.9	$0	0	($35)	0.4	($70)	0.6	LS33: Loomes and Sugden
54	($52.50)	($42.00)	25%	$0	0.25	($35)	0	($70)	0.75	$0	0.1	($35)	0.6	($70)	0.3	LS34: Loomes and Sugden
55	($52.50)	($35.00)	50%	$0	0.25	($35)	0	($70)	0.75	$0	0	($35)	1	($70)	0	LS35: Loomes and Sugden
56	($42.00)	($35.00)	20%	$0	0.1	($35)	0.6	($70)	0.3	$0	0	($35)	1	($70)	0	LS36: Loomes and Sugden
57	($28.00)	($21.00)	33%	$0	0.5	($35)	0.2	($70)	0.3	$0	0.4	($35)	0.6	($70)	0	LS37: Loomes and Sugden
58	($31.50)	($21.00)	50%	$0	0.55	($35)	0	($70)	0.45	$0	0.4	($35)	0.6	($70)	0	LS38: Loomes and Sugden
59	($31.50)	($28.00)	13%	$0	0.55	($35)	0	($70)	0.45	$0	0.5	($35)	0.2	($70)	0.3	LS39: Loomes and Sugden
60	($21.00)	($14.00)	50%	$0	0.7	($35)	0	($70)	0.3	$0	0.6	($35)	0.4	($70)	0	LS40: Loomes and Sugden
61	$59.50	$51.80	15%	($35)	0.1	($21)	0	$70	0.9	($35)	0	($21)	0.2	$70	0.8	LS9: Loomes and Sugden
62	$17.50	($13.30)	-232%	($35)	0.5	($21)	0	$70	0.5	($35)	0.1	($21)	0.8	$70	0.1	LS10: Loomes and Sugden
63	$17.50	($21.00)	-183%	($35)	0.5	($21)	0	$70	0.5	($35)	0	($21)	1	$70	0	LS11: Loomes and Sugden
64	($13.30)	($21.00)	-37%	($35)	0.1	($21)	0.8	$70	0.1	($35)	0	($21)	1	$70	0	LS12: Loomes and Sugden
65	($3.50)	($18.90)	-81%	($35)	0.7	($21)	0	$70	0.3	($35)	0.5	($21)	0.4	$70	0.1	LS13: Loomes and Sugden
66	($3.50)	($26.60)	-87%	($35)	0.7	($21)	0	$70	0.3	($35)	0.4	($21)	0.6	$70	0	LS14: Loomes and Sugden
67	($18.90)	($26.60)	-29%	($35)	0.5	($21)	0.4	$70	0.1	($35)	0.4	($21)	0.6	$70	0	LS15: Loomes and Sugden
68	($24.50)	($32.20)	-24%	($35)	0.9	($21)	0	$70	0.1	($35)	0.8	($21)	0.2	$70	0	LS16: Loomes and Sugden
69	$59.50	$33.60	77%	($35)	0.1	($21)	0	$70	0.9	($35)	0	($21)	0.4	$70	0.6	LS33: Loomes and Sugden
70	$43.75	$4.90	793%	($35)	0.25	($21)	0	$70	0.75	($35)	0.1	($21)	0.6	$70	0.3	LS34: Loomes and Sugden
71	$43.75	($21.00)	-308%	($35)	0.25	($21)	0	$70	0.75	($35)	0	($21)	1	$70	0	LS35: Loomes and Sugden
72	$4.90	($21.00)	-123%	($35)	0.1	($21)	0.6	$70	0.3	($35)	0	($21)	1	$70	0	LS36: Loomes and Sugden
73	($0.70)	($26.60)	-97%	($35)	0.5	($21)	0.2	$70	0.3	($35)	0.4	($21)	0.6	$70	0	LS37: Loomes and Sugden
74	$12.25	($26.60)	-146%	($35)	0.55	($21)	0	$70	0.45	($35)	0.4	($21)	0.6	$70	0	LS38: Loomes and Sugden
75	$12.25	($0.70)	-1850%	($35)	0.55	($21)	0	$70	0.45	($35)	0.5	($21)	0.2	$70	0.3	LS39: Loomes and Sugden
76	($3.50)	($29.40)	-88%	($35)	0.7	($21)	0	$70	0.3	($35)	0.6	($21)	0.4	$70	0	LS40: Loomes and Sugden
77	$52.50	$50.00	5%	$10	0.15	$20	0	$60	0.85	$10	0	$20	0.25	$60	0.75	LS1: Loomes and Sugden
78	$45.00	$42.50	6%	$10	0.3	$20	0	$60	0.7	$10	0.15	$20	0.25	$60	0.6	LS2: Loomes and Sugden
79	$45.00	$40.00	13%	$10	0.3	$20	0	$60	0.7	$10	0	$20	0.5	$60	0.5	LS3: Loomes and Sugden
80	$42.50	$40.00	6%	$10	0.15	$20	0.25	$60	0.6	$10	0	$20	0.5	$60	0.5	LS4: Loomes and Sugden
81	$22.50	$20.00	13%	$10	0.15	$20	0.75	$60	0.1	$10	0	$20	1	$60	0	LS5: Loomes and Sugden
82	$30.00	$20.00	50%	$10	0.6	$20	0	$60	0.4	$10	0	$20	1	$60	0	LS6: Loomes and Sugden
83	$30.00	$22.50	33%	$10	0.6	$20	0	$60	0.4	$10	0.15	$20	0.75	$60	0.1	LS7: Loomes and Sugden
84	$15.00	$12.50	20%	$10	0.9	$20	0	$60	0.1	$10	0.75	$20	0.25	$60	0	LS8: Loomes and Sugden
85	$50.00	$47.50	5%	$5	0.1	$25	0	$55	0.9	$5	0	$25	0.25	$55	0.75	LS17: Loomes and Sugden

86	$35.00	$27.50	27%	$5	0.4	$25	0	$55	0.6	$5	0.1	$25	0.75	$55	0.15	LS18: Loomes and Sugden
87	$35.00	$25.00	40%	$5	0.4	$25	0	$55	0.6	$5	0	$25	1	$55	0	LS19: Loomes and Sugden
88	$27.50	$25.00	10%	$5	0.1	$25	0.75	$55	0.15	$5	0	$25	1	$55	0	LS20: Loomes and Sugden
89	$20.00	$17.50	14%	$5	0.7	$25	0	$55	0.3	$5	0.6	$25	0.25	$55	0.15	LS21: Loomes and Sugden
90	$20.00	$15.00	33%	$5	0.7	$25	0	$55	0.3	$5	0.5	$25	0.5	$55	0	LS22: Loomes and Sugden
91	$17.50	$15.00	17%	$5	0.6	$25	0.25	$55	0.15	$5	0.5	$25	0.5	$55	0	LS23: Loomes and Sugden
92	$12.50	$10.00	25%	$5	0.85	$25	0	$55	0.15	$5	0.75	$25	0.25	$55	0	LS24: Loomes and Sugden
93	$42.00	$40.50	4%	$15	0.1	$30	0	$45	0.9	$15	0	$30	0.3	$45	0.7	LS25: Loomes and Sugden
94	$33.00	$30.00	10%	$15	0.4	$30	0	$45	0.6	$15	0.2	$30	0.6	$45	0.2	LS26: Loomes and Sugden
95	$33.00	$28.50	16%	$15	0.4	$30	0	$45	0.6	$15	0.1	$30	0.9	$45	0	LS27: Loomes and Sugden
96	$30.00	$28.50	5%	$15	0.2	$30	0.6	$45	0.2	$15	0.1	$30	0.9	$45	0	LS28: Loomes and Sugden
97	$27.00	$25.50	6%	$15	0.6	$30	0	$45	0.4	$15	0.5	$30	0.3	$45	0.2	LS29: Loomes and Sugden
98	$27.00	$24.00	13%	$15	0.6	$30	0	$45	0.4	$15	0.4	$30	0.6	$45	0	LS30: Loomes and Sugden
99	$25.50	$24.00	6%	$15	0.5	$30	0.3	$45	0.2	$15	0.4	$30	0.6	$45	0	LS31: Loomes and Sugden
100	$21.00	$19.50	8%	$15	0.8	$30	0	$45	0.2	$15	0.7	$30	0.3	$45	0	LS32: Loomes and Sugden

Table A2: Battery of 100 Lottery Tasks in Choices Made by MBA Students

Task	EV left	EV right	EV ratio	Left $ 1	Left p 1	Left $ 2	Left p 2	Left $ 3	Left p 3	Right $ 1	Right p 1	Right $ 2	Right p 2	Right $ 3	Right p 3	Notes
1	$5.00	$6.95	-0.2806	$0	0	$5	1	$0	0	$0	0.01	$5	0.89	$25	0.1	Allais - lower stakes
2	$0.55	$2.50	-0.78	$0	0.89	$5	0.11	$0	0	$0	0.9	$5	0	$25	0.1	Allais - lower stakes
3	$100.00	$139.00	-0.2806	$0	0	$100	1	$0	0	$0	0.01	$100	0.89	$500	0.1	Allais - higher stakes
4	$11.00	$50.00	-0.78	$0	0.89	$100	0.11	$0	0	$0	0.9	$100	0	$500	0.1	Allais - higher stakes
5	$425.00	$437.50	-0.029	$0	0.15	$250	0	$500	0.85	$0	0	$250	0.25	$500	0.75	LS1: Loomes and Sugden
6	$350.00	$362.50	-0.034	$0	0.3	$250	0	$500	0.7	$0	0.15	$250	0.25	$500	0.6	LS2: Loomes and Sugden
7	$350.00	$375.00	-0.067	$0	0.3	$250	0	$500	0.7	$0	0	$250	0.5	$500	0.5	LS3: Loomes and Sugden
8	$362.50	$375.00	-0.033	$0	0.15	$250	0.25	$500	0.6	$0	0	$250	0.5	$500	0.5	LS4: Loomes and Sugden
9	$237.50	$250.00	-0.05	$0	0.15	$250	0.75	$500	0.1	$0	0	$250	1	$500	0	LS5: Loomes and Sugden
10	$200.00	$250.00	-0.2	$0	0.6	$250	0	$500	0.4	$0	0	$250	1	$500	0	LS6: Loomes and Sugden
11	$200.00	$237.50	-0.1579	$0	0.6	$250	0	$500	0.4	$0	0.15	$250	0.75	$500	0.1	LS7: Loomes and Sugden
12	$50.00	$62.50	-0.2	$0	0.9	$250	0	$500	0.1	$0	0.75	$250	0.25	$500	0	LS8: Loomes and Sugden
13	$450.00	$450.00	0	$0	0.1	$250	0	$500	0.9	$0	0	$250	0.2	$500	0.8	LS9: Loomes and Sugden
14	$250.00	$250.00	0	$0	0.5	$250	0	$500	0.5	$0	0.1	$250	0.8	$500	0.1	LS10: Loomes and Sugden
15	$250.00	$250.00	0	$0	0.5	$250	0	$500	0.5	$0	0	$250	1	$500	0	LS11: Loomes and Sugden
16	$250.00	$250.00	0	$0	0.1	$250	0.8	$500	0.1	$0	0	$250	1	$500	0	LS12: Loomes and Sugden
17	$150.00	$150.00	0	$0	0.7	$250	0	$500	0.3	$0	0.5	$250	0.4	$500	0.1	LS13: Loomes and Sugden
18	$150.00	$150.00	0	$0	0.7	$250	0	$500	0.3	$0	0.4	$250	0.6	$500	0	LS14: Loomes and Sugden
19	$150.00	$150.00	0	$0	0.5	$250	0.4	$500	0.1	$0	0.4	$250	0.6	$500	0	LS15: Loomes and Sugden
20	$50.00	$50.00	0	$0	0.9	$250	0	$500	0.1	$0	0.8	$250	0.2	$500	0	LS16: Loomes and Sugden
21	$450.00	$437.50	0.0286	$0	0.1	$250	0	$500	0.9	$0	0	$250	0.25	$500	0.75	LS17: Loomes and Sugden
22	$300.00	$262.50	0.14286	$0	0.4	$250	0	$500	0.6	$0	0.1	$250	0.75	$500	0.15	LS18: Loomes and Sugden
23	$300.00	$250.00	0.2	$0	0.4	$250	0	$500	0.6	$0	0	$250	1	$500	0	LS19: Loomes and Sugden
24	$262.50	$250.00	0.05	$0	0.1	$250	0.75	$500	0.15	$0	0	$250	1	$500	0	LS20: Loomes and Sugden

25	$150.00	$137.50	0.0909	$0	0.7	$250	0	$500	0.3	$0	0.6	$250	0.25	$500	0.15	LS21: Loomes and Sugden
26	$150.00	$125.00	0.2	$0	0.7	$250	0	$500	0.3	$0	0.5	$250	0.5	$500	0	LS22: Loomes and Sugden
27	$137.50	$125.00	0.1	$0	0.6	$250	0.25	$500	0.15	$0	0.5	$250	0.5	$500	0	LS23: Loomes and Sugden
28	$75.00	$62.50	0.2	$0	0.85	$250	0	$500	0.15	$0	0.75	$250	0.25	$500	0	LS24: Loomes and Sugden
29	$450.00	$425.00	0.0588	$0	0.1	$250	0	$500	0.9	$0	0	$250	0.3	$500	0.7	LS25: Loomes and Sugden
30	$300.00	$250.00	0.2	$0	0.4	$250	0	$500	0.6	$0	0.2	$250	0.6	$500	0.2	LS26: Loomes and Sugden
31	$300.00	$225.00	0.33333	$0	0.4	$250	0	$500	0.6	$0	0.1	$250	0.9	$500	0	LS27: Loomes and Sugden
32	$250.00	$225.00	0.11111	$0	0.2	$250	0.6	$500	0.2	$0	0.1	$250	0.9	$500	0	LS28: Loomes and Sugden
33	$200.00	$175.00	0.14286	$0	0.6	$250	0	$500	0.4	$0	0.5	$250	0.3	$500	0.2	LS29: Loomes and Sugden
34	$200.00	$150.00	0.33333	$0	0.6	$250	0	$500	0.4	$0	0.4	$250	0.6	$500	0	LS30: Loomes and Sugden
35	$175.00	$150.00	0.16667	$0	0.5	$250	0.3	$500	0.2	$0	0.4	$250	0.6	$500	0	LS31: Loomes and Sugden
36	$100.00	$75.00	0.33333	$0	0.8	$250	0	$500	0.2	$0	0.7	$250	0.3	$500	0	LS32: Loomes and Sugden
37	$450.00	$400.00	0.125	$0	0.1	$250	0	$500	0.9	$0	0	$250	0.4	$500	0.6	LS33: Loomes and Sugden
38	$375.00	$300.00	0.25	$0	0.25	$250	0	$500	0.75	$0	0.1	$250	0.6	$500	0.3	LS34: Loomes and Sugden
39	$375.00	$250.00	0.5	$0	0.25	$250	0	$500	0.75	$0	0	$250	1	$500	0	LS35: Loomes and Sugden
40	$300.00	$250.00	0.2	$0	0.1	$250	0.6	$500	0.3	$0	0	$250	1	$500	0	LS36: Loomes and Sugden
41	$200.00	$150.00	0.33333	$0	0.5	$250	0.2	$500	0.3	$0	0.4	$250	0.6	$500	0	LS37: Loomes and Sugden
42	$225.00	$150.00	0.5	$0	0.55	$250	0	$500	0.45	$0	0.4	$250	0.6	$500	0	LS38: Loomes and Sugden
43	$225.00	$200.00	0.125	$0	0.55	$250	0	$500	0.45	$0	0.5	$250	0.2	$500	0.3	LS39: Loomes and Sugden
44	$150.00	$100.00	0.5	$0	0.7	$250	0	$500	0.3	$0	0.6	$250	0.4	$500	0	LS40: Loomes and Sugden
45	($450.00)	($450.00)	0	$0	0.1	($250)	0	($500)	0.9	$0	0	($250)	0.2	($500)	0.8	LS9: Loomes and Sugden
46	($250.00)	($250.00)	0	$0	0.5	($250)	0	($500)	0.5	$0	0.1	($250)	0.8	($500)	0.1	LS10: Loomes and Sugden
47	($250.00)	($250.00)	0	$0	0.5	($250)	0	($500)	0.5	$0	0	($250)	1	($500)	0	LS11: Loomes and Sugden
48	($250.00)	($250.00)	0	$0	0.1	($250)	0.8	($500)	0.1	$0	0	($250)	1	($500)	0	LS12: Loomes and Sugden
49	($150.00)	($150.00)	0	$0	0.7	($250)	0	($500)	0.3	$0	0.5	($250)	0.4	($500)	0.1	LS13: Loomes and Sugden
50	($150.00)	($150.00)	0	$0	0.7	($250)	0	($500)	0.3	$0	0.4	($250)	0.6	($500)	0	LS14: Loomes and Sugden
51	($150.00)	($150.00)	0	$0	0.5	($250)	0.4	($500)	0.1	$0	0.4	($250)	0.6	($500)	0	LS15: Loomes and Sugden
52	($50.00)	($50.00)	0	$0	0.9	($250)	0	($500)	0.1	$0	0.8	($250)	0.2	($500)	0	LS16: Loomes and Sugden
53	($450.00)	($400.00)	0.125	$0	0.1	($250)	0	($500)	0.9	$0	0	($250)	0.4	($500)	0.6	LS33: Loomes and Sugden
54	($375.00)	($300.00)	0.25	$0	0.25	($250)	0	($500)	0.75	$0	0.1	($250)	0.6	($500)	0.3	LS34: Loomes and Sugden
55	($375.00)	($250.00)	0.5	$0	0.25	($250)	0	($500)	0.75	$0	0	($250)	1	($500)	0	LS35: Loomes and Sugden
56	($300.00)	($250.00)	0.2	$0	0.1	($250)	0.6	($500)	0.3	$0	0	($250)	1	($500)	0	LS36: Loomes and Sugden
57	($200.00)	($150.00)	0.33333	$0	0.5	($250)	0.2	($500)	0.3	$0	0.4	($250)	0.6	($500)	0	LS37: Loomes and Sugden
58	($225.00)	($150.00)	0.5	$0	0.55	($250)	0	($500)	0.45	$0	0.4	($250)	0.6	($500)	0	LS38: Loomes and Sugden
59	($225.00)	($200.00)	0.125	$0	0.55	($250)	0	($500)	0.45	$0	0.5	($250)	0.2	($500)	0.3	LS39: Loomes and Sugden
60	($150.00)	($100.00)	0.5	$0	0.7	($250)	0	($500)	0.3	$0	0.6	($250)	0.4	($500)	0	LS40: Loomes and Sugden
61	$425.00	$370.00	0.14865	($250)	0.1	($150)	0	$500	0.9	($250)	0	($150)	0.2	$500	0.8	LS9: Loomes and Sugden
62	$125.00	($95.00)	-2.3158	($250)	0.5	($150)	0	$500	0.5	($250)	0.1	($150)	0.8	$500	0.1	LS10: Loomes and Sugden
63	$125.00	($150.00)	-1.8333	($250)	0.5	($150)	0	$500	0.5	($250)	0	($150)	1	$500	0	LS11: Loomes and Sugden
64	($95.00)	($150.00)	-0.3667	($250)	0.1	($150)	0.8	$500	0.1	($250)	0	($150)	1	$500	0	LS12: Loomes and Sugden
65	($25.00)	($135.00)	-0.8148	($250)	0.7	($150)	0	$500	0.3	($250)	0.5	($150)	0.4	$500	0.1	LS13: Loomes and Sugden
66	($25.00)	($190.00)	-0.8684	($250)	0.7	($150)	0	$500	0.3	($250)	0.4	($150)	0.6	$500	0	LS14: Loomes and Sugden
67	($135.00)	($190.00)	-0.2895	($250)	0.5	($150)	0.4	$500	0.1	($250)	0.4	($150)	0.6	$500	0	LS15: Loomes and Sugden
68	($175.00)	($230.00)	-0.2391	($250)	0.9	($150)	0	$500	0.1	($250)	0.8	($150)	0.2	$500	0	LS16: Loomes and Sugden
69	$425.00	$240.00	0.77083	($250)	0.1	($150)	0	$500	0.9	($250)	0	($150)	0.4	$500	0.6	LS33: Loomes and Sugden

70	$312.50	$35.00	7.92857	($250)	0.25	($150)	0	$500	0.75	($250)	0.1	($150)	0.6	$500	0.3	LS34: Loomes and Sugden
71	$312.50	($150.00)	-3.0833	($250)	0.25	($150)	0	$500	0.75	($250)	0	($150)	1	$500	0	LS35: Loomes and Sugden
72	$35.00	($150.00)	-1.2333	($250)	0.1	($150)	0.6	$500	0.3	($250)	0	($150)	1	$500	0	LS36: Loomes and Sugden
73	($5.00)	($190.00)	-0.9737	($250)	0.5	($150)	0.2	$500	0.3	($250)	0.4	($150)	0.6	$500	0	LS37: Loomes and Sugden
74	$87.50	($190.00)	-1.4605	($250)	0.55	($150)	0	$500	0.45	($250)	0.4	($150)	0.6	$500	0	LS38: Loomes and Sugden
75	$87.50	($5.00)	-18.5	($250)	0.55	($150)	0	$500	0.45	($250)	0.5	($150)	0.2	$500	0.3	LS39: Loomes and Sugden
76	($25.00)	($210.00)	-0.881	($250)	0.7	($150)	0	$500	0.3	($250)	0.6	($150)	0.4	$500	0	LS40: Loomes and Sugden
77	$52.50	$50.00	0.05	$10	0.15	$20	0	$60	0.85	$10	0	$20	0.25	$60	0.75	LS1: Loomes and Sugden
78	$45.00	$42.50	0.0588	$10	0.3	$20	0	$60	0.7	$10	0.15	$20	0.25	$60	0.6	LS2: Loomes and Sugden
79	$45.00	$40.00	0.125	$10	0.3	$20	0	$60	0.7	$10	0	$20	0.5	$60	0.5	LS3: Loomes and Sugden
80	$42.50	$40.00	0.0625	$10	0.15	$20	0.25	$60	0.6	$10	0	$20	0.5	$60	0.5	LS4: Loomes and Sugden
81	$22.50	$20.00	0.125	$10	0.15	$20	0.75	$60	0.1	$10	0	$20	1	$60	0	LS5: Loomes and Sugden
82	$30.00	$20.00	0.5	$10	0.6	$20	0	$60	0.4	$10	0	$20	1	$60	0	LS6: Loomes and Sugden
83	$30.00	$22.50	0.33333	$10	0.6	$20	0	$60	0.4	$10	0.15	$20	0.75	$60	0.1	LS7: Loomes and Sugden
84	$15.00	$12.50	0.2	$10	0.9	$20	0	$60	0.1	$10	0.75	$20	0.25	$60	0	LS8: Loomes and Sugden
85	$50.00	$47.50	0.0526	$5	0.1	$25	0	$55	0.9	$5	0	$25	0.25	$55	0.75	LS17: Loomes and Sugden
86	$35.00	$27.50	0.27273	$5	0.4	$25	0	$55	0.6	$5	0.1	$25	0.75	$55	0.15	LS18: Loomes and Sugden
87	$35.00	$25.00	0.4	$5	0.4	$25	0	$55	0.6	$5	0	$25	1	$55	0	LS19: Loomes and Sugden
88	$27.50	$25.00	0.1	$5	0.1	$25	0.75	$55	0.15	$5	0	$25	1	$55	0	LS20: Loomes and Sugden
89	$20.00	$17.50	0.14286	$5	0.7	$25	0	$55	0.3	$5	0.6	$25	0.25	$55	0.15	LS21: Loomes and Sugden
90	$20.00	$15.00	0.33333	$5	0.7	$25	0	$55	0.3	$5	0.5	$25	0.5	$55	0	LS22: Loomes and Sugden
91	$17.50	$15.00	0.16667	$5	0.6	$25	0.25	$55	0.15	$5	0.5	$25	0.5	$55	0	LS23: Loomes and Sugden
92	$12.50	$10.00	0.25	$5	0.85	$25	0	$55	0.15	$5	0.75	$25	0.25	$55	0	LS24: Loomes and Sugden
93	$42.00	$40.50	0.037	$15	0.1	$30	0	$45	0.9	$15	0	$30	0.3	$45	0.7	LS25: Loomes and Sugden
94	$33.00	$30.00	0.1	$15	0.4	$30	0	$45	0.6	$15	0.2	$30	0.6	$45	0.2	LS26: Loomes and Sugden
95	$33.00	$28.50	0.15789	$15	0.4	$30	0	$45	0.6	$15	0.1	$30	0.9	$45	0	LS27: Loomes and Sugden
96	$30.00	$28.50	0.0526	$15	0.2	$30	0.6	$45	0.2	$15	0.1	$30	0.9	$45	0	LS28: Loomes and Sugden
97	$27.00	$25.50	0.0588	$15	0.6	$30	0	$45	0.4	$15	0.5	$30	0.3	$45	0.2	LS29: Loomes and Sugden
98	$27.00	$24.00	0.125	$15	0.6	$30	0	$45	0.4	$15	0.4	$30	0.6	$45	0	LS30: Loomes and Sugden
99	$25.50	$24.00	0.0625	$15	0.5	$30	0.3	$45	0.2	$15	0.4	$30	0.6	$45	0	LS31: Loomes and Sugden
100	$21.00	$19.50	0.0769	$15	0.8	$30	0	$45	0.2	$15	0.7	$30	0.3	$45	0	LS32: Loomes and Sugden

Appendix B: Numerical Examples of Decision Weights (NOT FOR PUBLICATION)

To understand the mechanics of evaluating lotteries using RDU and CPT it is useful to see worked numerical examples. Although this is purely a pedagogic exercise, in our experience many users of RDU and CPT are not familiar with these mechanics, and they are critical to the correct application of these models. Even the best pedagogic source available, Wakker [2010], leaves many worked examples as exercises, and many of the examples are correctly contrived to make a special pedagogic point. The most general source, actually, is an online computer program on Peter Wakker's home page that calculates values for CPT using the Inverse-S probability weighting function and up to four outcomes:

http://people.few.eur.nl/wakker/miscella/calculate.cpt.kobb/index.htm

We use this program to generate some examples to illustrate the logic of decision weights under CPT.

The building block for understanding the construction of decision weights under CPT, for the general case of a mixed-frame lottery, is the construction of decision weights for gains under RDU. We provide one detailed example there, and then examine the CPT extension.

B.1 Rank-Dependent Decision Weights

Assume a simple power probability weighting function $\omega(p) = p^\gamma$ and let $\gamma = 1.25$. To see the pure effect of probability weighting, assume $U(x) = x$ for $x \geq 0$. Start with a two-prize lottery, then consider three-prizes and four-prizes to see the general logic.

In the two-prize case, let y be the smaller prize and Y be the larger prize, so $Y > y \geq 0$. Again, to see the pure effect of probability weighting, assume objective probabilities $p(y) = p(Y) = ½$. The first step is to get the decision weight of the largest prize. This uses the answer to the question, "what is the probability of getting at *least* Y?"[35] This is obviously ½, so we then calculate the decision weight using the probability weighting function as $\omega(½) = (½)^{\gamma} = 0.42$. To keep notation for probability weights and decision weights similar but distinct, denote the *decision* weight for Y as w(Y). Then we have w(Y) = 0.42.

The second step for the two-prize case is to give the other, smaller prize y the residual weight. This uses the answer to the question, "what is the probability of getting at *least* y?" Since one always gets at least y, the answer is obviously 1. Since $\omega(1) = 1$ for any of the popular probability weighting functions,[36] we can attribute the decision weight $\omega(1) - \omega(½) = 1 - 0.42 = 0.58$ to the prize y. Another way to see the same thing is to directly calculate the decision weight for the smallest prize to ensure that the decision weights sum to 1, so that the decision weight w(y) is calculated as 1-w(Y) = 1 - 0.42 = 0.58. The two-prize case actually makes it harder to see the rank-dependent logic than when we examine the three-prize or four-prize case, but can be seen in retrospect as a special case.

With these two decision weights in place, the RDU evaluation of the lottery is $0.42 \times U(Y) + 0.58 \times U(y)$, or 0.42Y + 0.58y given our simplifying assumption of a linear utility function. Inspection of this RDU evaluation, and viewing the decision weights as if they were probabilities, shows why the RDU evaluation has to be less than the Expected Value (EV) of the lottery using the true probabilities, since that is 0.5Y + 0.5y. The RDU evaluation puts more weight on the worst prize, and greater weight on the better prize, so it has to have a CE that is less than the EV (this last step is helped by the fact that U(x) = x, of course). Hence probability weighting in this case generates a CE that is less than the EV, and hence a risk premium.

[35] This expression leads to what Wakker [2010; §7.6] usefully calls the "gain-rank." The "loss-rank" would be based on the answer to the question, "what is the probability of getting Y or *less*?" Loss-ranks were popular with some of the earlier studies in rank-dependent utility.

[36] The prominent exception is the probability weighting function suggested by Kahneman and Tversky [1979], which had interior discontinuities at p=0 and p=1.

However, the two-prize case collapses the essential logic of the RDU model. Consider a three-prize case in which we use the same probability weighting functions and utility functions, but have three prizes, y, Y and **Y**, where **Y**>Y>y, and p(y) = p(Y) = p(**Y**) = ⅓.

The decision weight for **Y** is evaluated first, and uses the answer to the question, "what is the probability of getting at least **Y**?" The answer is ⅓, so the decision weight for **Y** is then directly evaluated as $w(\mathbf{Y}) = \omega(⅓) = (⅓)^{\gamma} = 0.25$.

The decision weight for Y is evaluated next, and uses the answer to the more interesting question, "what is the probability of getting at *least* Y?" This is $p(Y) + p(\mathbf{Y}) = ⅓ + ⅓ = ⅔$, so the probability weight is $\omega(⅔) = (⅔)^{\gamma} = 0.60$. But the only part of this probability weight that is to be attributed solely to Y is the part that is not already attributed to **Y**, hence the decision weight for Y is $\omega(⅔) - \omega(⅓) = \omega(Y) - \omega(\mathbf{Y}) = 0.60 - 0.25 = 0.35$. This intermediate step shows the rank-dependent logic in the clearest fashion. One could equally talk about *cumulative* probability weights, rather than just probability weights, but the logic is simple enough when one thinks of the question being asked "psychologically" and the partial attribution to Y that flows from it. In the two-prize case this partial attribution is skipped over.

The decision weight for y is again evaluated residually, as in the two-prize case. We can either see this by evaluating $\omega(1) - \omega(⅔) = 1 - 0.60 = 0.40$, or by evaluating $1 - w(Y) - w(\mathbf{Y}) = 1 - 0.35 - 0.25 = 0.40$.

The general logic may now be stated in words as follows:

- Rank the prizes from best to worst.
- Use the probability weighting function to calculate the probability of getting at least the prize in question.
- Then assign the decision weight for the best prize directly as the weighted probability of that prize.

- For each of the intermediate prizes in declining order, assign the decision weight using the weighted cumulative probability for that prize less the decision weights for better prizes (or, equivalently, the weighted cumulative probability for the immediately better prize).
- For the worst prize the decision weight is the residual decision weight to ensure that the decision weights sum to 1.

The key is to view the decision weights as the *incremental* decision weight attributable to that prize.

Table B1 collects these steps for each of the examples, and adds a four prize example. From a programming perspective, these calculations are tedious but not difficult as long as one can assume that prizes are rank-ordered as they are evaluated. Our computer code in *Stata* allows for up to four prizes, which spans most applications in laboratory or field settings, and is of course applicable for lotteries with any number of prizes up to four. The logic can be easily extended to more prizes.

Figure B1 illustrates these calculations using the power probability weighting function. The dashed line in the left panel displays the probability weighting function $\omega(p) = p^{\gamma} = p^{1.25}$, with the vertical axis showing underweighting of the objective probabilities displayed on the bottom axis. The implications for decision weights are then shown in the right panel, for the two-prize, three-prize and four-prize cases. In the right panel the bottom axis shows prizes ranked from worst to best, so one immediately identifies the "probability pessimism" at work with this probability weighting function. Values of $\gamma < 1$ generate overweighting of the objective probabilities and "probability optimism," as one might expect.

Figure B2 shows the effects of using the "inverse-S" probability weighting function $\omega(p) = p^{\gamma} / (p^{\gamma} + (1-p)^{\gamma})^{1/\gamma}$ for $\gamma = 0.65$. This function exhibits inverse-S probability weighting (optimism for small p, and pessimism for large p) for $\gamma<1$, and S-shaped probability weighting (pessimism for small p, and optimism for large p) for $\gamma>1$. Although one observes a wide range of values of γ in careful applied work, for many CPT advocates the qualitative assumption that $\gamma<1$ is often regarded as a critical component of CPT.

B.2 Cumulative Prospect Theory Decision Weights

The calculation of decision weights for CPT builds on this RDU logic. Indeed, for lotteries that are in the gain frame, there is nothing to add. But for lotteries that are in the loss frame or mixed frame, one has to be careful in applying these procedures.

Loss Frame Decision Weights

Consider a lottery that pays -100, -50, -25 and 0 with equal probability ¼ and EV of -43.75. Assume a Power probability weighting function, since this specification allows us to look at the effects of underweighting and overweighting without worrying about whether the probabilities are small or large. Start with the overweighting case, in which γ- is 0.5. The tabulations are shown in panel A of Table B2. The first point of difference with the way in which decision weights were calculated for gains in Table B1 is that we have listed the probabilities from Worst to Best, rather than Best to Worst. In fact, that is the only point of difference, apart from using the weighting function with γ- = 0.5. This is an application of the notion of loss-ranks, stressed by Wakker [2010; §7.6] in an explication of the history of RDU thought, and then in Wakker [2010; §9.1] when introducing the sign-dependence of CPT. Gain-ranks are used for gains under CPT, and loss-ranks are used for losses under CPT. This point is so important that it is worth being verbose and restate the RDU logic using the language of gain-ranks, so that the parallel with loss-ranks becomes evidence.

For gains, in panel C of Table B1, we rank the *best* prize as #1, the second best prize as #2, the third best prize as #3, and the *worst* prize as #3. The probability whose outcome has *gain*-rank 1 is assigned a decision weight that is equal to the weighted probability for that outcome (0.18). Then the probability whose outcome has *gain*-rank 2 is assigned a decision weight equal to the weighted cumulative probability of getting that outcome (0.42) minus the weighted cumulative probability of getting the outcome with *gain*-rank 1 (0.18, so the decision weight is 0.24 = 0.42-0.18). And so on for the remaining outcomes.

Now turn to panel A of Table B2, with losses. We rank the ***worst*** prize as #1, the second worst prize as #2, the third worst prize as #3, and the *best* prize as #3. The probability whose outcome has *loss*-rank 1 is assigned a decision weight that is equal to the weighted probability for that outcome (0.50). Then the probability whose outcome has *loss*-rank 2 is assigned a decision weight equal to the weighted cumulative probability of getting that outcome (0.71) minus the weighted cumulative probability of getting the outcome with loss-rank 1 (0.50, so the decision weight is 0.21 = 0.71-0.50). And so on for the remaining outcomes.

To drive home the parallel nature of the calculations, once the shift from gain-ranks to loss-ranks has been made, in panel B of Table B2 we consider an underweighting Power probability weighting with γ- = 1.25, exactly the same function that was used in panel C of Table B1. Apart from the listing of probabilities from worst outcome to best outcome, the calculations are identical!

Putting aside the mechanics of calculating these decision weights, focus on the effect of overweighting and underweighting on the final decision weights, recalling of course that the underlying objective probabilities were each ¼. In the overweighting case the decisions weights put greater weight on the worst outcomes compared to the best outcomes, so if the utility function was linear, it is apparent that the CE would be lower than the EV by construction, implying a positive risk premium and, *ceteris paribus*, risk aversion. In the underweighting case the decisions weights put greater weight on the best outcomes compared to the worst outcomes, so if the utility function was linear the CE would be higher than the EV, implying a negative risk premium and, *ceteris paribus*, risk seeking. So we end up with the reverse effect of overweighting and underweighting in terms of risk aversion compared to the gain frame.

Mixed Frame Decision Weights

The logic here is to initially parse the mixed frame lottery in a gain lottery and a loss lottery, evaluate each of those two parsed lotteries while ensuring that each has residual weight on the "zero outcome," and then add the parsed evaluations. The "zero outcome" here is in quotation marks because it refers to the assumed reference point, which need not be zero in any currency units, although it often is. In our analysis the "zero outcome" is in fact the endowment given to subjects, and hence it embodies the assumption that subjects fail to locally asset integrate the endowment with the framed prizes. The logic of the example below is general, but one needs to keep this distinction in mind when generalizing CPT to other reference points.

We employ the following notation, found in many CPT studies. Rank order the n outcomes so that

$$x_1 \geq \dots \geq x_k \geq 0 \geq x_{k+1} \geq \dots \geq x_n$$

so that the k^{th} and $k+1^{th}$ outcomes mark the dividing line between gains and losses. Then define the mixed-frame lottery as

$$P: (x_1, p_1; \dots ; x_k, p_k; 0, p_0; x_{k+1}, p_{k+1}; \dots ; x_n, p_n)$$

Note that it is quite possible that $p_0 = 0$ in the specification of P. The mixed-frame lottery is then parsed into a gain frame component and a loss frame component. The gain frame component is defined as

$$P+: (x_1, p_1; \dots ; x_k, p_k; 0, p_0 + p_{k+1} + \dots + p_n),$$

where the "zero outcome" is assigned all of the probability mass for *losses*, as well as any probability mass originally assigned to p_0. The loss frame component is defined as

$$P\text{-}: (0, p_1 + ... + p_k + p_0; x_{k+1}, p_{k+1}; ... ; x_n, p_n),$$

where the "zero outcome" is assigned all of the probability mass for *gains*, as well as any probability mass originally assigned to p_0. One then evaluates the cumulative prospective utility of P+ and P-, which we can denote CPU(P+) and CPU(P-), and then these are literally added together to get the cumulative prospective utility of the mixed-frame lottery P:

$$CPU(P) = CPU(P+) + CPU(P-)$$

Thus if we know how to evaluate the RDU decision weights for P+ using gain-ranks, and the decision weights for loss frame prospects such as P- using loss-ranks, we can generate the decision weights for all types of lotteries using CPT.[37]

To take a textbook example, consider Exercise 9.34 from Wakker [2010; p.257]. In this case $u(x) = x^\alpha$ for $x \geq 0$ and $u(x) = x^\beta$ for $x<0$, with $\alpha = \beta = 0.5$, $\omega(p+) = (p+)^2$ for gains, $\omega(p-) = p-$ for losses, and $\lambda = 2$. Let the mixed frame gamble be

P: (9, 0.1; 4, 0.4; -4, 0.4, -9, 0.1)

so the parsed gain frame component is

P+: (9, 0.1; 4, 0.5; 0, 0.5)

and the parsed loss frame component is

P-: (0, 0.5; -4, 0.4, -9, 0.1).

Tables B3 and B4 show the detailed application of this parsing process.

To take a full-blooded example, consider the default lottery on the web page referred to earlier. This specification uses the Inverse-S probability weighting function, with parameters $\gamma+ = 0.61$ and $\gamma- = 0.69$ in terms of our notation. There are two negative prizes, and two positive prizes in the lottery, and equal probability ¼ for each outcome. Figure B3 shows an image of the output of

[37] Wakker [2010; p. 255, 261] discusses the implications of the decisions weights for mixed frame lotteries summing to more than 1. There are none.

the web page after evaluating this lottery, with the decision weights at the bottom. Table B5 spells out the calculations in the manner in which we have been presenting them, ending up with exactly the same answers apart from trivial rounding errors. We explicitly list the "fake 0 reference point" that is added to P+ and P-, although these are arithmetically irrelevant. The symbol ✂ means that the numerical value in that cell is cut out, since it is not needed to identify the final decision weights.

Table B1: Tabulations of RDU Examples

Prize	Probability	Cumulative Probability	Weighted Cumulative Probability	Decision Weight
		A. Two Prizes		
Y	0.5	0.5	$0.42 = 0.5^{1.25}$	0.42
y < Y	0.5	1	$1 = 1^{1.25}$	0.58 = 1 - 0.42
		B. Three Prizes		
Y	0.33	0.33	$0.25 = 0.33^{1.25}$	0.25
Y < **Y**	0.33	0.67	$0.60 = 0.67^{1.25}$	0.35 = 0.60 - 0.25
y < Y < **Y**	0.33	1	$1 = 1^{1.25}$	0.40 = 1 - 0.60 = 1 - 0.35 - 0.25
		C. Four Prizes		
Best	0.25	0.25	$0.18 = 0.25^{1.25}$	0.18
2nd Best	0.25	0.5	$0.42 = 0.50^{1.25}$	0.24 = 0.42 - 0.18
3rd Best	0.25	0.75	$0.70 = 0.75^{1.25}$	0.28 = 0.70 - 0.42 = 1 - 0.24 - 0.18
Worst	0.25	1	$1^{1.25}$	0.30 = 1 - 0.70 = 1 - 0.28 - 0.24 - 0.18

Figure B1: Power Probability Weighting and Implied Decision Weights for Gains

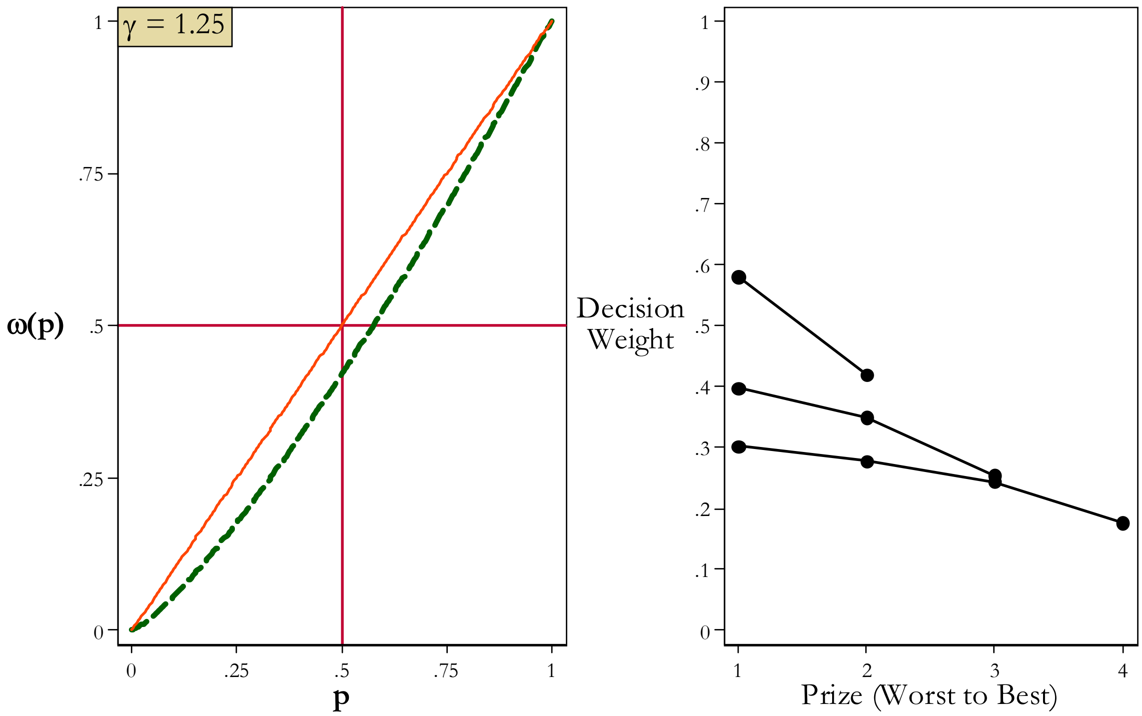


Figure B2: Inverse-S Probability Weighting and Implied Decision Weights for Gains

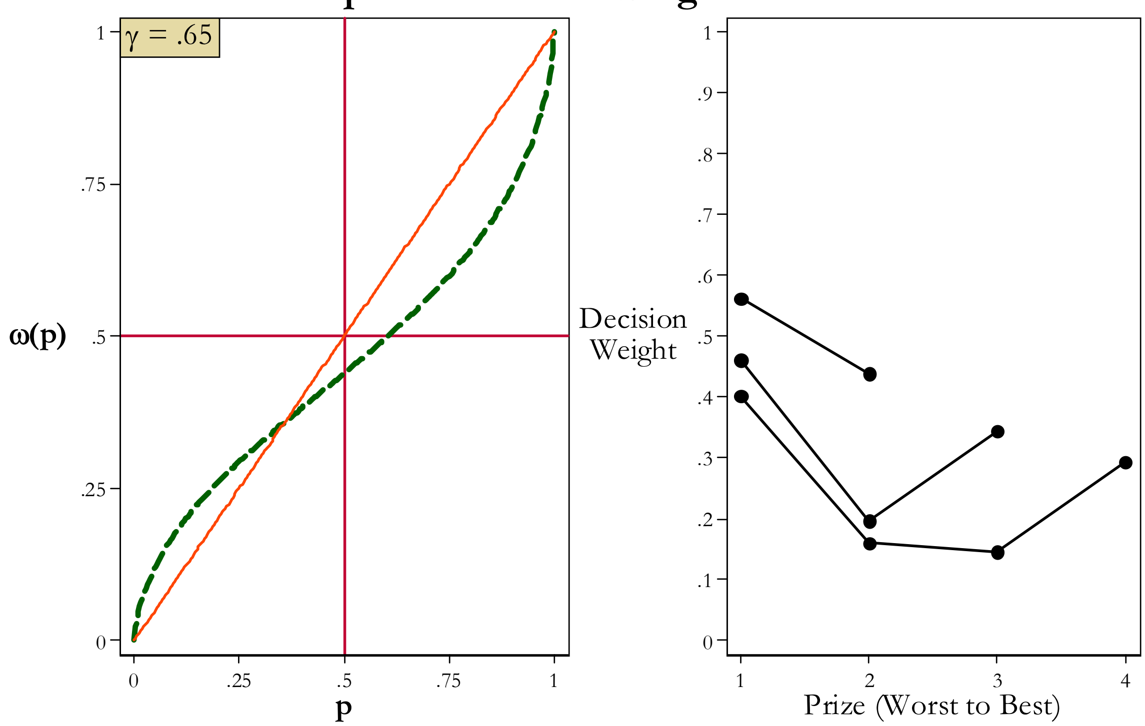


Table B2: Tabulations of CPT Loss Frame Examples

Prize	Probability	Cumulative Probability	Weighted Cumulative Probability	Decision Weight
		A. Overweighting		
Worst	0.25	0.25	$0.50 = 0.25^{0.5}$	0.5
2nd Worst	0.25	0.5	$0.71 = 0.50^{0.5}$	0.21 = 0.71 - 0.50
3rd Worst	0.25	0.75	$0.87 = 0.75^{0.5}$	0.16 = 0.87 - 0.71
Best	0.25	1	$1^{0.5}$	0.13 = 1 - 0.87
		B. Underweighting		
Worst	0.25	0.25	$0.18 = 0.25^{1.25}$	0.18
2nd Worst	0.25	0.5	$0.42 = 0.50^{1.25}$	0.24 = 0.42 - 0.18
3rd Worst	0.25	0.75	$0.70 = 0.75^{1.25}$	0.28 = 0.70 - 0.42
Best	0.25	1	$1^{1.25}$	0.30 = 1 - 0.70

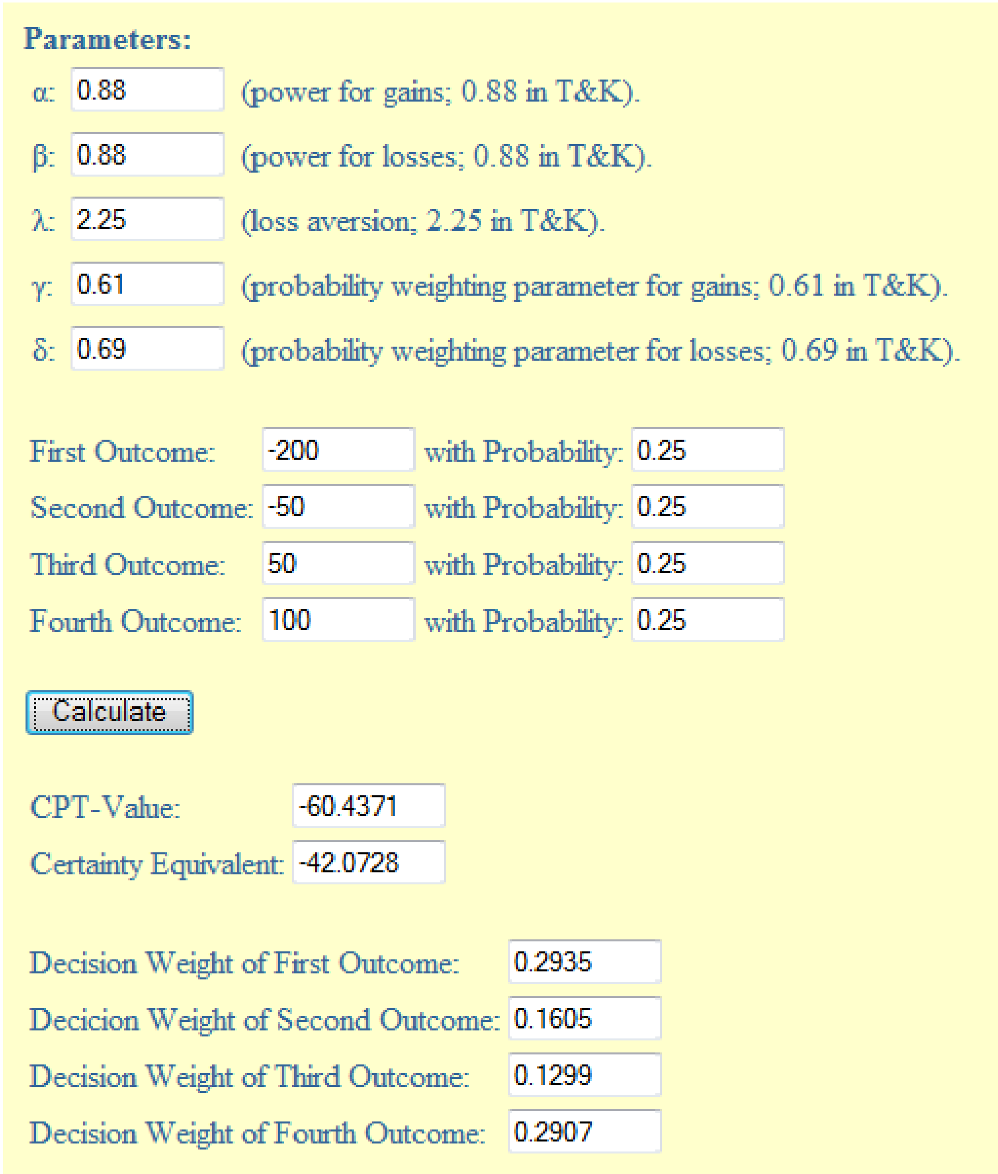
Table B3: Initial Tabulation of CPT Example

x	p	u(x)	-λu(-x)	x+ or x-	Σ p	ω (Σ p)	w(x)
			I. Evaluation of P+				
-9	0.1		-6 = -2×√9	0	1	1	0.19
-4	0.4		-4 = -2×√4	0	0.9	0.81	0.56
4	0.4	2 = √4		4	0.5	0.25	0.24
9	0.1	3 = √9		9	0.1	0.01	0.01
			II. Evaluation of P-				
-9	0.1		-6 = -2×√9	-9	0.1	0.1	0.1
-4	0.4		-4 = -2×√4	-4	0.5	0.5	0.4
4	0.4	2 = √4		0	0.9	0.9	0.4
9	0.1	3 = √9		0	1	1	0.1

Table B4: Final Tabulation of CPT Example

x	p	U(x) = u(x)	U(x) = -λ×u(-x)	w(x) if x≥0	w(x) if x<0	CPU(x) = w(x)×U(x)	EU(x) = p×u(x)
-9	0.1		-6		0.1	-0.6	-0.3
-4	0.4		-4		0.4	-1.6	-0.8
4	0.4	2		0.24		0.48	0.8
9	0.1	3		0.01		0.03	0.3
Sum						-1.69	0

Figure B3: Default Example from Wakker Web Page



Parameters:

α: 0.88 (power for gains; 0.88 in T&K).

β: 0.88 (power for losses; 0.88 in T&K).

λ: 2.25 (loss aversion; 2.25 in T&K).

γ: 0.61 (probability weighting parameter for gains; 0.61 in T&K).

δ: 0.69 (probability weighting parameter for losses; 0.69 in T&K).

First Outcome: -200 with Probability: 0.25

Second Outcome: -50 with Probability: 0.25

Third Outcome: 50 with Probability: 0.25

Fourth Outcome: 100 with Probability: 0.25

Calculate

CPT-Value: -60.4371

Certainty Equivalent: -42.0728

Decision Weight of First Outcome: 0.2935

Decicion Weight of Second Outcome: 0.1605

Decision Weight of Third Outcome: 0.1299

Decision Weight of Fourth Outcome: 0.2907

Table B5: Tabulations of CPT Mixed Frame Examples from Wakker Home Page

Prize	Probability	Cumulative Probability	Weighted Cumulative Probability	Decision Weight
		A. Gain Frame		
Best (200)	0.25	0.25	0.29	0.29
2^{nd} Best (50)	0.25	0.5	0.42	0.13 = 0.42 - 0.29
3^{rd} Best (Fake 0)	0.25	0.75	✂	✂
Worst (Fake 0)	0.25	1	✂	✂
		B. Loss Frame		
Worst (-200)	0.25	0.25	0.29	0.29
2^{nd} Worst (-50)	0.25	0.5	0.45	0.16 = 0.45 - 0.29
3^{rd} Worst (Fake 0)	0.25	0.75	✂	✂
Best (Fake 0)	0.25	1	✂	✂

Appendix C: Main Instructions (NOT FOR PUBLICATION)

Choices Over Risky Prospects

This is a task where you will choose between prospects with varying prizes and chances of winning. You will be presented with a series of pairs of prospects where you will choose one of them. There are 100 pairs in the series. For each pair of prospects, you should choose the prospect you prefer to play. You will actually get the chance to play one of the prospects you choose, and you will be paid according to the outcome of that prospect, so you should think carefully about which prospect you prefer.

Here is an example of what the computer display of such a pair of prospects will look like.

The outcome of the prospects will be determined by the draw of a random number between 1 and 100. Each number between, and including, 1 and 100 is equally likely to occur. In fact, you will be able to draw the number yourself using two 10-sided dice.

You will be told your cash endowment for each lottery at the top of the lottery. In this example it is $35, so any earnings would be added to or subtracted from this endowment. The endowment may change from choice to choice, so be sure to pay attention to it. The endowment you are shown only applies for that choice.

In the above example the left prospect pays twenty-five dollars ($25) if the number drawn is between 1 and 5, and pays negative five dollars ($-5) if the number is between 6 and 55, and pays negative thirty-five dollars ($-35) if the number is between 56 and 100. The blue color in the pie chart corresponds to 5% of the area and illustrates the chances that the number drawn will be between 1 and 5 and your prize will be $25. The orange area in the pie chart corresponds to 50% of the area and illustrates the chances that the number drawn will be between 6 and 55 and your prize will be $-5. The green area in the pie chart corresponds to 45% of the area and illustrates the chances that the number drawn will be between 56 and 100. When you select the lottery to be played out the computer will tell you what die throws translate into what prize.

Now look at the pie in the chart on the right. It pays twenty-five dollars ($25) if the number drawn is between 1 and 15, negative five dollars ($-5) if the number is between 16 and 25, and negative thirty-five dollars ($-35) if the number is between 26 and 100. As with the prospect on the left, the pie slices represent the fraction of the possible numbers which yield each payoff. For example, the size of the $25 pie slice is 15% of the total pie.

Even though the screen says that you might win a negative amount, this is actually a loss to be deducted from your endowment. So if you "win" $-5, your earnings would be $30 = $35 - $5.

Each pair of prospects is shown on a separate screen on the computer. On each screen, you should indicate which prospect you prefer to play by clicking on one of the buttons beneath the prospects.

After you have worked through all of the pairs of prospects, raise your hand and an experimenter will come over. You will then roll two ten-sided die to determine which pair of the 100 prospects you chose will be played out:. Since there is a chance that any of your 100 choices could be played out for real, you should approach each pair of prospects as if it is the one that you will play out. Finally, you will roll the two ten-sided dice again to determine the outcome of the prospect you chose.

For instance, suppose you picked the prospect on the right in the above example. If the random number was 7, you would win $25 in addition to your endowment; if it was 93, you would lose $35 from your endowment. If you picked the prospect on the left and drew the number 7, you would lose $5 from your endowment; if it was 93, you would again lose $35 from your endowment.

Therefore, your payoff is determined by three things:

- by which prospect you selected, the left or the right, for each of these 100 pairs;
- by which prospect pair is chosen to be played out in the series of 100 such pairs using the two ten-sided die; and
- by the outcome of that prospect when you roll the two 10-sided dice.

Which prospects you prefer is a matter of personal taste. The people next to you may be presented with different prospects, and may have different preferences, so their responses should not matter to you. Please work silently, and make your choices by thinking carefully about each prospect.

All payoffs are in cash, and are in addition to the $7.50 show-up fee that you receive just for being here, as well as any other earnings in other tasks.

Appendix D: Additional Instructions (NOT FOR PUBLICATION)

E1. Introductory Text

Today's Experiment

Thank you for participating in this experiment. Please turn off your cell phones and any other non-health related devices (electronic or otherwise) that you may have. The use of such devices is prohibited while in the lab.

In this experiment, you will participate in more than one decision-making task. You will have the opportunity to earn money in the first task. In the second task, you stand a chance of winning or losing money.

After the first task, which will earn you money, you will make 100 decisions in the second task. One of your second task decisions will be selected at random to be played out for cash. Since each of those second task decisions is equally likely to be played for cash, you should carefully consider each decision.

You have already earned $10 for your participation in today's experiment. You will NOT be asked to risk your $10 participation fee. Any other money that you earn in today's experiment will depend on your choices, and also on chance. However, you will not leave the experiment with any earned money (other than the $10 participation fee) unless you complete the entire experiment today.

If you wish to withdraw at this time or at any time before the end of experiment, you may do so and keep only your $10 participation fee. Please initial this form to indicate that you understand the requirements above.

E2. Quiz for Earned Endowment

The following questions test your knowledge of current events, American history, and geography. Please indicate the correct answer in the blank beside each question. You will be paid based on the number of questions you answer correctly.

If you answer 8 or more questions correctly, you will be paid $80.
If you answer 7 or fewer questions correctly, you will be paid $40.

______1. The current Secretary of State is

a. Joe Biden
b. Timothy Geithner
c. John Kerry
d. Hillary Clinton

______2. The winner of the 2013 Superbowl was

a. New York Giants
b. San Fransisco 49ers
c. Green Bay Packers
d. Baltimore Ravens

______3. Which of the following states borders the Gulf of Mexico?

a. California
b. Texas
c. Maine
d. North Carolina

______4. Who was the last President to die in office?

a. John Kennedy
b. Bill Clinton
c. Gerald Ford
d. Ronald Reagan

______5. What is the capital of Arkansas?

a. Pierre
b. Sacramento
c. Albany
d. Little Rock

______6. Which of the following was one of the first 13 colonies?

a. Montana
b. Virginia
c. Louisiana
d. Texas

______7. Who is the host of American Idol?

a. Howie Mandel
b. Regis Philben
c. Jeff Probst
d. Ryan Seacrest

______8. Which of the following toys was named for a U.S. President?

a. Jacks
b. Raggedy Andy
c. Marco Polo
d. Teddy bear

______9. "Only you can prevent wild fires." is the slogan of

a. Toucan Sam
b. Polly the Parrot
c. Woodsy the Owl
d. Smokey the Bear

______10. Which of the following was an ally of the United States in World War II?

a. Germany
b. Switzerland
c. Italy
d. Great Britain

______11. Which of the following is a movie set in a future where the Capitol selects a boy and girl from the twelve districts to fight to the death on live television?

a. 21 Jump Street
b. Ted
c. The Hunger Games
d. Safe House

______12. Which television network carries the Real Housewives?

a. Bravo
b. PBS
c. HBO
d. MTV

______13. “Grey’s Anatomy” is a television series centered around

a. a carwash
b. a hospital
c. a baseball team
d. hotel maid service

______14. Who is credited with inventing the light bulb?

a. Eli Whitney
b. Oprah Winfrey
c. Thomas Edison
d. Enrico Marconi

______15. “First in Flight” is the slogan of which of the following states?

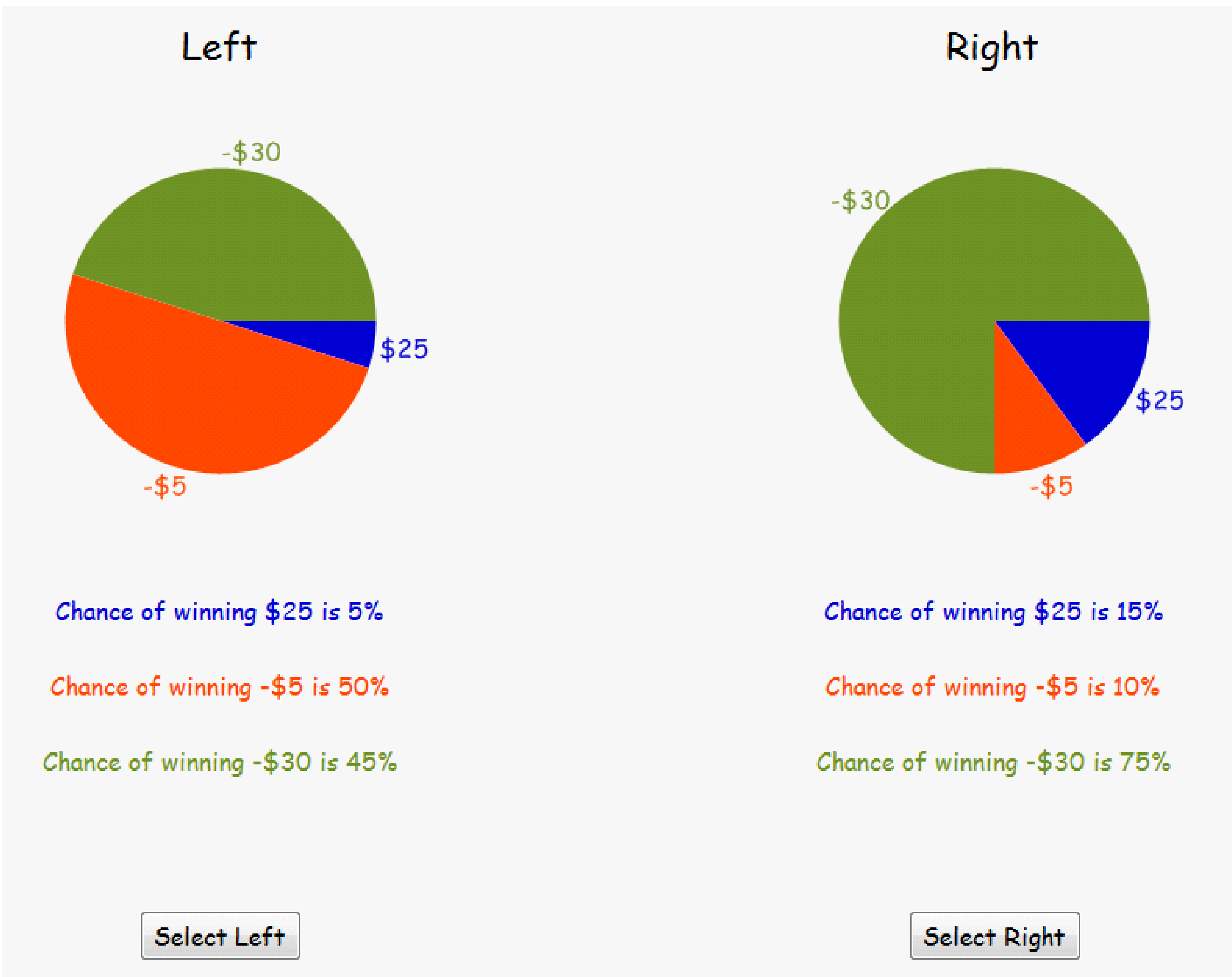
a. Texas
b. Montana
c. Maine
d. North Carolina

E3. Choice Task Instructions

Choices Over Risky Prospects

This is a task where you will choose between prospects with varying prizes and chances of winning. You will be presented with a series of pairs of prospects where you will choose one of them. There are 100 pairs in the series. For each pair of prospects, you should choose the prospect you prefer to play. You will actually get the chance to play one of the prospects you choose, and you will be paid according to the outcome of that prospect, so you should think carefully about which prospect you prefer.

Here is an example of what the computer display of such a pair of prospects will look like.



The outcome of the prospects will be determined by the draw of a random number between 1 and 100. Each number between, and including, 1 and 100 is equally likely to occur. In fact, you will be able to draw the number yourself using two 10-sided dice.

Any losses from this task will be taken from your initial earnings sitting on your desk.

In the above example the left prospect pays twenty-five dollars ($25) if the number drawn is between 1 and 5, and pays negative five dollars ($-5) if the number is between 6 and 55, and pays negative thirty dollars ($-30) if the number is between 56 and 100. The blue color in the pie chart corresponds to 5% of the area and illustrates the chances that the number drawn will be between 1 and 5 and your prize will be $25. The orange area in the pie chart corresponds to 50% of the area and illustrates the chances that the number drawn will be between 6 and 55 and your prize will be $-5. The green area in the pie chart corresponds to 45% of the area and illustrates the chances that the number drawn will be between 56 and 100 and your prize will be $-30. When you select the lottery to be played out the computer will tell you what die throws translate into what prize.

Now look at the pie in the chart on the right. It pays twenty-five dollars ($25) if the number drawn is between 1 and 15, negative five dollars ($-5) if the number is between 16 and 25, and negative thirty dollars ($-30) if the number is between 26 and 100. As with the prospect on the left, the pie slices represent the fraction of the possible numbers which yield each payoff. For example, the size of the $25 pie slice is 15% of the total pie.

Even though the screen says that you might win a negative amount, this is actually a loss to be deducted from your quiz earnings. So if you "win" $-5 and you earned $50 from the quiz, your earnings would be $45 = $50 - $5.

Each pair of prospects is shown on a separate screen on the computer. On each screen, you should indicate which prospect you prefer to play by clicking on one of the buttons beneath the prospects.

Since there is a chance that any of your choices could be played out for real, you should approach each pair of prospects as if it is the one that you will play out. When one of your choices from this task is selected to be played out, you will roll the two ten-sided dice to determine the outcome of the prospect you chose.

For instance, suppose you picked the prospect on the right in the above example. If the random number was 7, you would win $25 in addition to your endowment; if it was 93, you would lose $30 from your endowment. If you picked the prospect on the left and drew the number 7, you would lose $5 from your endowment; if it was 93, you would lose $30 from your endowment.

Which prospects you prefer is a matter of personal taste. The people next to you may be presented with different prospects, and may have different preferences, so their responses should not matter to you. Please work silently, and make your choices by thinking carefully about each prospect.

All payoffs are in cash, and are in addition to the $10 show-up payment that you receive just for being here. Any losses will be deducted from your initial earnings. You will never lose your show-up payment.