Econometrica Supplementary Material

SUPPLEMENT TO "ON THE EMPIRICAL VALIDITY OF CUMULATIVE PROSPECT THEORY: EXPERIMENTAL EVIDENCE OF RANK-INDEPENDENT PROBABILITY WEIGHTING"

(*Econometrica*, Vol. 88, No. 4, July 2020, 1363–1409)

B. DOUGLAS BERNHEIM
Department of Economics, Stanford University and NBER

CHARLES SPRENGER
Rady School of Management and Department of Economics, University of California San Diego

APPENDIX A: RE-EXAMINATION OF PRIOR PROSPECT THEORY ELICITATION DATA

EXPERIMENTS DESIGNED TO ELICIT PROSPECT THEORY PARAMETERS such as Tversky and Kahneman (1992), Tversky and Fox (1995), and Gonzalez and Wu (1999) generally have subjects provide certainty equivalents for binary lotteries. For example, Tversky and Kahneman (1992) elicited certainty equivalents for a 10%, 50%, and 90% chance of receiving \$50 with the alternative being zero, and also elicited certainty equivalents for a 10%, 50%, and 90% chance of receiving \$50 with the alternative being \$100.

One may wish to use such data to examine whether a given probability of receiving \$50 is weighted differently depending on its rank. Note that binary lotteries generally do not permit meaningful tests of the core axioms of comonotonic and non-comonotonic independence in the vein of Wu (1994) and Wakker, Erev, and Weber (1994) because two binary lotteries with a common outcome will have a dominance relation. Nonetheless, parametric estimates using binary lottery data could, in principle, support an interpretation of rank dependence in probability weights.

For lotteries with a p-probability of receiving \$50 and an alternative of \$0, Tversky and Kahneman (1992) reported median certainty equivalents for $p \in \{0.1, 0.5, 0.9\}$ of $\{\$9, \$21, \$37\}$. For lotteries with a p-probability of receiving \$50 and an alternative of \$100, Tversky and Kahneman (1992) reported median certainty equivalents for $p \in \{0.1, 0.5, 0.9\}$ of $\{\$83, \$71, \$59\}$.

Using these two data sets, one could estimate probability weighting and curvature under the null hypothesis of rank independence and then test that null. That is, for each lottery, one assumes the indifference condition

$$C = u^{-1}\big(\pi(p)u(50) + \pi(1-p)u(X)\big) + \epsilon$$

is satisfied, where X is either \$0 or \$100 depending on the lottery in question.[1] Given the two-parameter model and nonlinear estimation techniques described in Section 4.1, with three observations we can estimate both the probability weighting parameter of $\pi(\cdot)$, γ, and the utility curvature parameter of $u(\cdot)$, α, with one degree of freedom in each series. Conducting such an exercise using the reported median data for lotteries between \$50 and \$0, we find $\gamma = 0.64$ and $\alpha = 0.98$. Conducting such an exercise using the reported

B. Douglas Bernheim: bernheim@stanford.edu
Charles Sprenger: csprenger@ucsd.edu

[1]Note that this formulation is not equivalent to that of Kahneman and Tversky (1979) for binary lotteries, which, by their equation (2), is rank-dependent.

 https://doi.org/10.3982/ECTA16646

median data for lotteries between \$50 and \$100, we find $\gamma = 0.55$ and $\alpha = 1.99$. Strictly speaking, these point estimates are inconsistent with the null hypothesis of rank independence. Consider a 90% chance of receiving \$50 when the alternative is \$0. With $\gamma = 0.64$, $\pi(0.9) = 0.74$. Now consider a 90% chance of receiving \$50 when the alternative is \$100. With $\gamma = 0.55$, $\pi(0.9) = 0.66$. Thus, the 90% chance of \$50 receives either 74% or 66% of the decision weight depending on whether the alternative is higher or lower than \$50. Setting aside the question of statistical precision, these estimates are inconsistent with the null hypothesis of rank independence.[2]

Exercises such as the one described above suffer from a fundamental identification problem. If one does not make specific functional assumptions about the shape of utility, the same data are reconcilable with rank independence. Let $w_H(p)$ and $w_L(p)$ represent the weight applied to a \$50 payoff when it is higher than the alternative (i.e., \$0) or lower than the alternative (i.e., \$100), respectively. The certainty equivalents for such prospects are

$$w_H(p)u(50) + w_H(1-p)u(0) = u(c_1),$$
$$w_L(p)u(50) + w_L(1-p)u(100) = u(c_2).$$

The weighting function is rank-independent if $w_H(p) = w_L(p) = w(p)$. In such a case,

$$w(1-p) = \frac{u(c_2) - u(c_1)}{u(100) - u(0)}.$$

Appropriate choice of utility function $u(\cdot)$ can rationalize the behavior c_1 and c_2 with a rank-independent weighting function. For example, focusing on Kahneman and Tversky's data for $p = 0.9$, rationalization requires

$$w(1-0.9) = \frac{u(59) - u(37)}{u(100) - u(0)}.$$

Thus, to rationalize all the data from Tversky and Kahneman (1992) with a rank-independent weighting function, one need only find $u(\cdot)$ and $w(\cdot)$ such that

$$w(0.1) = \frac{u(59) - u(37)}{u(100) - u(0)},$$
$$w(0.5) = \frac{u(71) - u(21)}{u(100) - u(0)},$$
$$w(0.9) = \frac{u(83) - u(9)}{u(100) - u(0)}.$$

This exercise demonstrates that interpreting data from binary lotteries as evidence for (or against) rank dependence is problematic. Different assumptions about the shape of utility can lead to qualitative differences in the extent of apparent rank dependence. One clear benefit of our proposed test of rank dependence is that, at its core, it is free from functional form assumptions both for the shape of utility and probability weighting.

[2]This conclusion is not altered (although the direction changes) if one imposes a common value of $\alpha = 0.98$ across the two data sets. The estimated γ for the alternative of \$100 becomes 0.70 and a 90% chance of \$50 receives a decision weight of 78%.

APPENDIX B: EXAMPLES OF CONFOUNDS AFFECTING EXISTING TESTS OF RANK DEPENDENCE

In Section 2.3, we explained that existing tests of rank dependence are difficult to interpret without a parametric model of noisy choice. In this appendix, we provide examples to illustrate the conceptual points made in the text.

As noted in Section 2.3, prior experiments in this domain have compared the frequencies of comonotonic independence (CI) and non-comonotonic independence (NCI) violations. One first elicits a binary preference between two comonotonic lotteries, S and R, that share a payoff event. One tests CI by replacing the shared payoff with another payoff that does not alter the ranking of outcomes, and eliciting preferences between the new options, S' and R'. One tests NCI by replacing the shared payoff with another payoff that does alter the ranking of outcomes, and eliciting preferences between the new options S'' and R''. For example, in one series of tasks, Wakker, Erev, and Weber (1994) considered the comonotonic lotteries

$$S = (\{0.55, 0.25, 0.2\}; \{0.5, 6.0, 7.0\}), \qquad R = (\{0.55, 0.25, 0.2\}; \{0.5, 4.5, 9.0\}).$$

They replaced the common 55% chance of 0.50 with 3.50 to construct

$$S' = (\{0.55, 0.25, 0.2\}; \{3.5, 6.0, 7.0\}), \qquad R' = (\{0.55, 0.25, 0.2\}; \{3.5, 4.5, 9.0\}),$$

which preserves the ranking. They replaced the common 55% chance of 3.50 with 6.50 to construct

$$S'' = (\{0.55, 0.25, 0.2\}; \{6.5, 6.0, 7.0\}), \qquad R'' = (\{0.55, 0.25, 0.2\}; \{6.5, 4.5, 9.0\}),$$

which alters the rankings. CPT requires a stable preference between (S, R) and (S', R'), an implication of CI, but permits preference reversals between (S', R') and (S'', R''), a failure of NCI.

Given that rank-dependent models permit violations of NCI, but not CI, some have used the relative frequency of CI and NCI violations in such environments as a measure of empirical support for rank dependence. The predominant finding is that decisionmakers violate both CI and NCI with high frequency, and at roughly the same rates.[3] Some interpret this finding as casting doubt on the validity of rank dependence.

Two features of these experiments preclude strong inferences and may have limited the impact of these works.

First, as explained in the text, the premise of the approach—that violation frequencies are necessarily higher for invalid axioms—is flawed. For reasonable models of noisy choice, noise-induced violations of independence are more likely to occur when the parameters of the choice tasks place the decisionmaker closer to the point of indifference. Existing approaches provide no way to ensure that the "distance from indifference" is held constant when comparing CI and NCI violations. Accordingly, one has no way of knowing whether the frequency of CI violations provides a valid benchmark for judging whether and to what extent the frequency of NCI violations is elevated. It is potentially an apples-to-oranges comparison.[4]

---

[3] Wakker, Erev, and Weber (1994) considered 12 CI tests and 6 NCI tests for each subject. The violation rates for both CI and NCI are around 40%. Wu (1994) presented similar tests and found CI violation rates of 47–50% and NCI violation rates of 38–50%.

[4] To make formal comparisons, Wakker, Erev, and Weber (1994) explicitly assumed that noise produces the same rate of violations for all choices. That assumption is obviously problematic, as one would expect violations

The following simple example starkly illustrates the problems resulting from this first point. We envision a CPT subject who obeys CI but not NCI. As noted above, we test CI by comparing choices between lotteries S and R with choices between lotteries S' and R'. Assume the subject has a "true" strict preference between S and R, and necessarily the same preference between S' and R', but because of (independent) noise chooses both S and S' with probability p. In that case, we will observe violations of CI with probability $2p(1-p)$. Likewise, we test NCI by comparing choices between S' and R' with choices between S'' and R''. Assume the resulting change in probability weighting yields a strong preference, so that S'' is chosen over R'' with probability 1. In that case, the frequency of observed NCI violations will be $1-p$. The difference between the frequency of NCI and CI violations is then $(1-p)(1-2p)$. A couple of observations follow. First, if the subject is initially close to indifference, so that p is close to 0.5, the observed differences in violation frequencies will be close to zero. Second, if $p > 0.5$, one will actually observe a higher frequency of violations for CI than for NCI, despite the fact that CPT is valid.

Second, even if one could control for "distance to indifference," existing approaches offer no basis for judging whether a given discrepancy between the frequencies of CI and NCI violations is large or small relative to the implications of a reasonably parameterized "noisy" CPT model. The following example illustrates how, even with constant "distance to indifference," one could find little or no difference between violation frequencies for CI and NCI, even though the rank-dependent formulation is correct. Assume in particular that, when confronted with a choice between two lotteries, the decisionmaker behaves according to the following noisy version of CPT: with probability p, she flips a coin; with probability $1-p$, she picks the best alternative according to a stable CPT objective function. Now suppose the experimental tasks are inadvertently chosen so that the typical subject is always far from indifference, with the unintended implication that rank reversals have no effect on the optimal choice according to the CPT representation. In that case, true rank dependence will not give rise to any NCI violations. Thus, the observed frequencies of CI and NCI violations will be identical $(p/2)$, even though CPT is the right theory, subject to noise.

Our illustrations are admittedly extreme. However, our point is general: without having a parameterized model of noisy choice and a method of gauging distance from indifference, there is simply no way to judge whether the discrepancy between the frequencies of CI and NCI violations is out of line with the implications of CPT.

APPENDIX C: EQUALIZING REDUCTIONS UNDER DIFFERENT REFERENCE POINT FORMULATIONS

C.1. *Fixed Referents*

This section investigates the predictions of CPT decisionmaking under alternative locations of an exogenous reference point. Under CPT, the decisionmaker is assumed to separate gains and losses and weight the corresponding probabilities separately. Gains are weighted according to the cumulative distribution beginning with the best possible outcome, while losses are weighted according to the decumulative distribution beginning with the worst possible outcome. CPT also allows for differences in the extent of probability weighting for gains and losses, $\pi_+(\cdot)$ and $\pi_-(\cdot)$, and the shape of utility for gains and losses, $u_+(\cdot)$ and $u_-(\cdot)$.

---

to be much more common for tasks that place the decisionmaker close to the point of indifference, which is what we assume for our next illustration.

In Section 2.2, we explained that CPT robustly implies a discontinuous change in decision weights when X crosses Y or Z, and that the percentage change in equalizing reduction robustly approximates the percentage change in relative decision weights. One complication noted in the text is that for non-infinitesimal values of m, $Y+m$, $Z-\underline{k}$, or $Z-\overline{k}$ may cross the reference point.

In order to examine the effect of crossing the reference point, Figure A1 provides simulations for $\overline{k}$, $\underline{k}$, and $\Delta\log(k)$ for $Z = \$18$, $Y = \$24$, $\underline{X} = 23$, and $\overline{X} = \$30$ at values of the reference point, $r \in (0, 40)$ for each of our probability vectors. Following Tversky and Kahneman (1992), we assume that gain and loss probability weighting functions are identical, $\pi_-(p) = \pi_+(p) = p^\gamma/(p^\gamma + (1-p)^\gamma)^{1/\gamma}$, with $\gamma = 0.61$. We also assume a piecewise linear formulation for loss averse utility such that $u_-(-x) = -\lambda u_+(x)$ with $u_+(x) = x$. The value of λ varies across rows. In addition to predicted behavior, we also provide estimates of $\Delta\log(w_Y/w_Z)$ for $\gamma = 0.61$ and the relevant probability vector for each condition.

Provided $r < Z-\underline{k}, Z-\overline{k}$ or $r > Y+m$, the values of $\Delta\log(k)$ closely approximate the change in weights $\Delta\log(w_Y/w_Z)$. Note, however, that because probability weighting is reference dependent, the relevant theoretical benchmark shifts from $\log(\pi(p+q) - \pi(p)) - \log(\pi(q))$ when $r < Z-\underline{k}, Z-\overline{k}$ to $\log(\pi(1-p) - \pi(1-p-q)) - \log(1 - \pi(1-q))$ when $r > Y+m$.

Figure A1 also illustrates two regions of transition. The first region encompasses $r \in (Z-\underline{k}, Z)$. In this region, log changes in behavior deviate from the theoretical benchmark. As r passes $Z-\underline{k}$, $\underline{k}$ is determined both by loss aversion, λ, and the weight attached to $Z-\underline{k}$ when it is considered a loss, $\pi(1-p-q)$. Once r passes $Z-\overline{k}$, the same is true of $\overline{k}$. When $Z-\underline{k}, Z-\overline{k} < r < Z$, $\Delta\log(k) \neq \log(\pi(p+q) - \pi(p)) - \log(\pi(q))$. However, the simple difference,

$$(\overline{k} - \underline{k})_{Z-\underline{k}, Z-\overline{k}<r<Z} = \frac{\pi(p+q) - \pi(p) - \pi(q)}{\lambda\pi(1-p-q)} m,$$

can be related to the prior difference when $r < Z-\underline{k}$,

$$(\overline{k} - \underline{k})_{r<Z-\underline{k}} = \frac{\pi(p+q) - \pi(p) - \pi(q)}{1 - \pi(p+q)} m.$$

Whether the difference $(\overline{k} - \underline{k})$, and hence $\Delta\log(k)$, grows or shrinks relative to this prior case depends on the value of λ and the difference between $\pi(1-p-q)$ and $1 - \pi(p+q)$. For our probability vectors, with $\gamma = 0.61$, $\pi(1-p-q) < 1 - \pi(p+q)$. As such, reference-dependent probability distortions, alone, would lead to larger values of $\Delta\log(k)$ in this region, and more apparent evidence of rank dependence. The top panel of Figure A1 illustrates this case with $\lambda = 1$. Values of $\lambda > 1$ counteract the force of probability weighting in this region. It must be noted, however, that even with substantial loss aversion of $\lambda = 2$, large negative values of $\Delta\log(k)$ are still predicted. When r passes Z, simulated behavior once again accords with the theoretical benchmark.

A second transition region arises for $r \in (Y, Y+m)$. Because Y and $Y+m$ are treated asymmetrically, $\underline{k}$ and $\overline{k}$ are functions both of reference-dependent probability

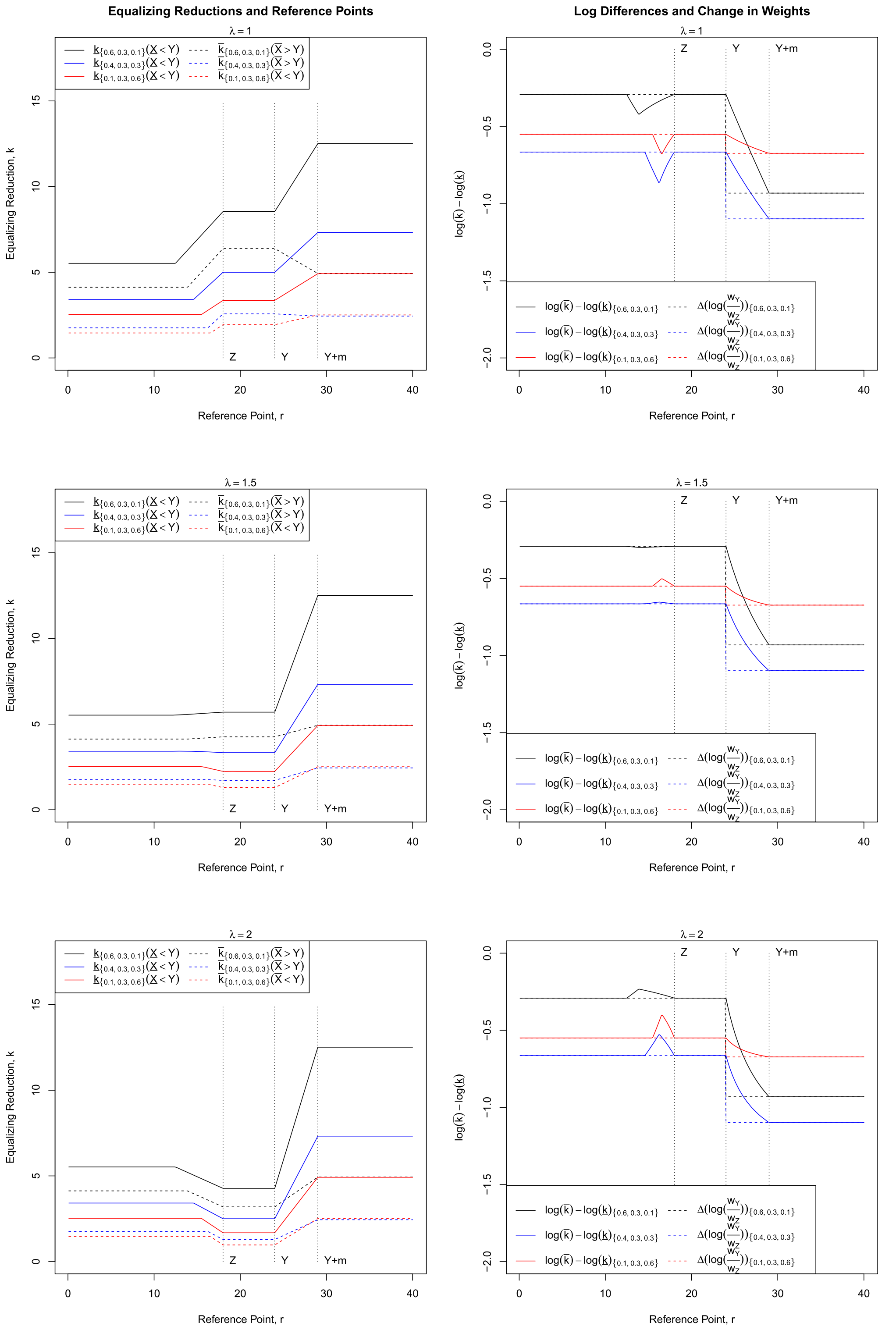

FIGURE A1.—Fixed referents and equalizing reductions.

distortions, and loss aversion. Specifically,

$$\underline{k} = \frac{\pi(q)}{\pi(1-p-q)\lambda}m + \frac{\left[\pi(q) - \big(1-\pi(1-q)\big)\lambda\right]}{\pi(1-p-q)\lambda}(Y-r),$$

$$\overline{k} = \frac{\big(\pi(p+q) - \pi(p)\big)}{\pi(1-p-q)\lambda}m + \frac{\left\{\left[\big(\pi(p+q)-\pi(p)\big)\right] - \left[\big(\pi(1-p) - \pi(1-p-q)\big)\lambda\right]\right\}}{\pi(1-p-q)\lambda}(Y-r),$$

in the transitional region. When $r \to Y$,

$$\Delta \log(k) \to \log\big(\pi(p+q) - \pi(p)\big) - \log\big(\pi(q)\big),$$

and when $r \to Y + m$,

$$\Delta \log(k) \to \log\big(\pi(1-p) - \pi(1-p-q)\big) - \log\big(1 - \pi(1-q)\big),$$

exactly the theoretical benchmarks at the region end-points. Though reference-dependent probability distortions determine the end-points of the transitional region, Figure A1 illustrates that the value of λ governs the speed of transition.

An interesting implication of CPT, which we mention in the main text, is that the relative decision weights on Y and Z do not just depend on their relationship to X, but also on the relationship of all three to the reference point. Examples of such effects are readily observed in Figure A1. As r passes the key points of Z and Y, equalizing reductions change abruptly, regardless of ranking information. This observation suggests a method for empirically identifying reference points: look for values of X, Y, Z at which the equalization reduction changes even though payoff ranks remain fixed.

These simulations show that CPT under standard parametric assumptions predicts sizable differences between $\underline{k}$ and $\overline{k}$ when the values of m and k are non-infinitesimal regardless of the location of the reference point, and that the percentage change in the equalizing reduction continues to approximate the percentage change in the relative decision weights outside of narrow regions of transition.

C.2. *Endogenous Referents*

Section 2.2 also provided a discussion of endogenous reference-points as in the models of disappointment aversion (DA) due to Bell (1985) and Loomes and Sugden (1986). In DA, the reference point is taken to be the EU certainty equivalent of the lottery in question, c. Here, we point out another feature of such models: even without rank-dependent probability weighting, these models imply the existence of a discontinuity in the equalizing reduction when X crosses the certainty equivalent, c. By varying X over a range that encompasses plausible values of c, one can therefore either identify the reference point or, in failing to find a discontinuity (as in our data), reject the theory.

Consider lottery L, which yields $X > Y > Z$ with corresponding probabilities p, q, $1-p-q$. Absent any additional probability weighting, the disappointment-averse representation is

$$U_{\text{DA}}(L) = pu(X|c) + qu(Y|c) + (1-p-q)u(Z|c),$$

where

$$c = v^{-1}\big(pv(X) + qv(Y) + (1 - p - q)v(Z)\big).$$

The reference-dependent utility is formalized as

$$u(x|r) = v(x) + \mu\big(v(x) - v(r)\big).$$

Assume a piecewise-linear gain-loss utility function,

$$\mu(y) = \begin{Bmatrix} \eta \cdot y & \text{if } y \geq 0, \\ \eta \cdot \lambda \cdot y & \text{if } y < 0, \end{Bmatrix}$$

where the parameter η captures sensitivity to gains and losses and λ represents the degree of loss aversion.[5] Note that this piecewise-linear form for reference dependence rules out a possibility discussed in Section 2.2: non-constancy of the X versus k schedule within regimes for which the ranks of X, Y, Z, and the reference point are fixed. Under this formulation, for $\overline{X}$ treated as a gain, $\overline{k}/m$ remains an approximation for the marginal rate of substitution between Y and Z:

$$\begin{aligned} &\text{MRS}_{YZ}(\overline{X} > Y, c) \\ &\quad = \big[\big(q + \eta q - \eta pq - \eta q^2 - \eta\lambda q(1 - p - q)\big) \\ &\qquad /\big((1 - p - q) + \eta\lambda(1 - p - q) - \eta p(1 - p - q) \\ &\qquad - \eta q(1 - p - q) - \eta\lambda(1 - p - q)^2\big)\big]\left[\frac{v'(Y)}{v'(Z)}\right] \\ &\quad \approx \frac{\overline{k}}{m}. \end{aligned}$$

If one lowers X to $\underline{X}$, but it remains treated as a gain relative to c, one predicts no change in equalizing reduction. However, if $\underline{X}$ is low enough to be considered a loss

[5]Whether Y is a loss or a gain depends on the exact values, probabilities, and shape of the utility function. Here, we analyze the case where Y is a gain and the addition of m and subtraction of k does not alter any gain-loss comparisons. In this case,

$$\begin{aligned} U_{\text{DA}}(L) = {} & \big[p + \eta p - \eta p^2 - \eta pq - \eta\lambda p(1 - p - q)\big]v(\overline{X}) \\ & + \big[q + \eta q - \eta pq - \eta q^2 - \eta\lambda q(1 - p - q)\big]v(Y) \\ & + \big[(1 - p - q) + \eta\lambda(1 - p - q) - \eta p(1 - p - q) - \eta q(1 - p - q) - \eta\lambda(1 - p - q)^2\big]v(Z), \end{aligned}$$

a formulation which 'weights' each outcome. If the addition of m to Y and subtraction of k from Z does not alter any gain-loss comparisons, the weights are the same for the equivalent lottery, L_e. As in our general formulation, the equalizing reduction captures the relative weights for outcomes Y and Z.

relative to c,[6] one finds

$$\begin{aligned}
&\mathrm{MRS}_{YZ}(\underline{X} < Y, c) \\
&\quad = \big[\big(q + \eta q - \eta\lambda pq - \eta q^2 - \eta\lambda q(1-p-q)\big) \\
&\qquad /\big((1-p-q) + \eta\lambda(1-p-q) - \eta\lambda p(1-p-q) \\
&\qquad - \eta q(1-p-q) - \eta\lambda(1-p-q)^2\big)\big]\left[\frac{v'(Y)}{v'(Z)}\right] \\
&\quad \approx \frac{k}{m}
\end{aligned}$$

and

$$\begin{aligned}
\Delta\log(k) = {} & \log\left(\frac{q + \eta q - \eta pq - \eta q^2 - \eta\lambda q(1-p-q)}{q + \eta q - \eta\lambda pq - \eta q^2 - \eta\lambda q(1-p-q)}\right) \\
& + \log\big(\big((1-p-q) + \eta\lambda(1-p-q) - \eta\lambda p(1-p-q) \\
& - \eta q(1-p-q) - \eta\lambda(1-p-q)^2\big) \\
& /\big((1-p-q) + \eta\lambda(1-p-q) - \eta p(1-p-q) \\
& - \eta q(1-p-q) - \eta\lambda(1-p-q)^2\big)\big) \\
\approx {} & \Delta\log(\mathrm{MRS}_{YZ}).
\end{aligned}$$

The value of X crossing the endogenous reference point of c leads to a discontinuity in the marginal rate of substitution and, hence, in equalizing reductions without explicit distortions of probabilities.

C.3. *Endogenous Reference Distributions*

Koszegi and Rabin (2006, 2007) (KR) built upon DA by assuming that the referent is dependent on the entire distribution of expected outcomes. An additional innovation of Koszegi and Rabin (2006, 2007) is a rational expectations equilibrium concept, the Unacclimating Personal Equilibrium (UPE). The objective of the UPE concept is to represent the notion that a rational individual will employ a reference distribution that coincides with the distribution of outcomes that will actually follow from her choices. The KR theory also features two refinements, Preferred Personal Equilibrium (PPE) and Choice-acclimating Personal Equilibrium (CPE).[7] We apply CPE when deriving the predictions

---

[6]For $\underline{X}$ low enough to be considered a loss, one arrives at

$$\begin{aligned}
U_{\mathrm{DA}}(L) = {} & \big[p + \eta\lambda p - \eta\lambda p^2 - \eta pq - \eta\lambda p(1-p-q)\big]v(\underline{X}) \\
& + \big[q + \eta q - \eta\lambda pq - \eta q^2 - \eta\lambda q(1-p-q)\big]v(Y) \\
& + \big[(1-p-q) + \eta\lambda(1-p-q) - \eta\lambda p(1-p-q) - \eta q(1-p-q) - \eta\lambda(1-p-q)^2\big]v(Z).
\end{aligned}$$

Note that the weights on both Y and Z have changed relative to the previous case. As before, the equalizing reduction summarizes these new relative weights.

[7]Both concepts maintain that the choice with the highest ex ante expected utility is selected. The operational distinction between the two concepts is that a CPE need not be a UPE, but a PPE must be a UPE.

of KR. That is, we assume the equalizing reduction corresponds to the point where the the decisionmaker switches from choosing L to L_e in CPE.

Let r represent a possible reference point drawn according to measure F. Let x be an outcome drawn according to the same measure F. Then the KR CPE utility formulation is the double integral

$$U(F|F) = \iint u(x|r)\,dF(r)\,dF(x)$$

with $u(x|r)$ as in DA. Under these preferences, the utility of lottery L, which yields $X > Y > Z$ with probabilities p, q, $(1-p-q)$, is

$$\begin{aligned} U_{\mathrm{KR}}(L|L) = {} & p\big(p\big[v(X)\big] + q\big[v(Y) + \eta\lambda\big(v(Y) - v(X)\big)\big] \\ & + (1-p-q)\big[v(Z) + \eta\lambda\big(v(Z) - v(X)\big)\big]\big) \\ & + q\big(p\big[v(X) + \eta\big(v(X) - v(Y)\big)\big] + q\big[v(Y)\big] \\ & + (1-p-q)\big[v(Z) + \eta\lambda\big(v(Z) - v(Y)\big)\big]\big) \\ & + (1-p-q)\big(p\big[v(X) + \eta\big(v(X) - v(Z)\big)\big] \\ & + q\big[v(Y) + \eta\big(v(Y) - v(Z)\big)\big] + (1-p-q)\big[v(Z)\big]\big). \end{aligned}$$

As with other models, $\overline{k}/m$ remains an approximation for the marginal rate of substitution between Y and Z, when $\overline{X} > Y$:

$$\begin{aligned} \mathrm{MRS}_{YZ}(\overline{X} > Y) = {} & \left[\frac{\big(q + pq\eta(\lambda-1) + q(1-p-q)\eta(1-\lambda)\big)}{\big((1-p-q) + p(1-p-q)\eta(\lambda-1) + q(1-p-q)\eta(\lambda-1)\big)}\right] \\ & \times \left[\frac{v'(Y)}{v'(Z)}\right] \\ \approx {} & \frac{\overline{k}}{m}. \end{aligned}$$

For $\underline{X} < Y$, the gain-loss comparisons are altered relative to the prior case, and

$$\begin{aligned} \mathrm{MRS}_{YZ}(\underline{X} < Y) = {} & \left[\frac{\big(q + pq\eta(1-\lambda) + q(1-p-q)\eta(1-\lambda)\big)}{\big((1-p-q) + p(1-p-q)\eta(\lambda-1) + q(1-p-q)\eta(\lambda-1)\big)}\right] \\ & \times \left[\frac{v'(Y)}{v'(Z)}\right] \\ \approx {} & \frac{\underline{k}}{m}. \end{aligned}$$

As X passes below Y, the marginal rate of substitution, and hence the equalizing reduction, changes discontinuously with

$$\Delta\log(k) = \log\left(\frac{\big(q + pq\eta(\lambda-1) + q(1-p-q)\eta(1-\lambda)\big)}{\big(q + pq\eta(1-\lambda) + q(1-p-q)\eta(1-\lambda)\big)}\right) \approx \Delta\log(\mathrm{MRS}_{YZ}).$$

TABLE AI
KOSZEGI–RABIN PREFERENCES[a]

$\{p, q, 1-p-q\}$	$\eta=1, \lambda=1.5$			$\eta=1, \lambda=2$		
	$\underline{k}$	$\overline{k}$	$\Delta\log(k)$	$\underline{k}$	$\overline{k}$	$\Delta\log(k)$
{0.6, 0.3, 0.1}	6.72	12.93	0.65	2.37	11.84	1.61
{0.4, 0.3, 0.3}	2.41	3.89	0.48	0.88	3.24	1.30
{0.1, 0.3, 0.6}	1.35	1.56	0.14	0.54	0.89	0.51

[a] Simulated values of $\underline{k}$ and $\overline{k}$ under Koszegi–Rabin preferences.

Even without explicit probability weighting, the KR theory carries implications of rank dependence and can also be tested by comparing equalizing reductions at different ranks.

To get a sense for magnitudes, Table AI simulates behavior under KR preferences in our experiment with $v(x) = x$, $\eta = 1$ and $\lambda = 1.5, 2$.[8] These simulations show that under the KR model, substantial discontinuities in equalizing reductions should be observed, in contrast to our findings.[9]

APPENDIX D: RANDOM CHOICE

In Section 5.4, we addressed the possibility that we detect no rank dependence because our subjects ignore the parameters of their decision tasks (either in general, or X in particular) and make their choices more or less randomly. In this appendix, we examine this possibility more formally by considering two explicit models. First, we consider individuals who choose randomly in each row of each equalizing reduction task. Such individuals would be expected to exhibit patterns of multiple switching many times in our experiment, which we do not observe. Standard practice in the experimental literature has been to take the first switch point as the relevant decision for such subjects. We reproduce our aggregate and individual graphs under this hypothesis in Figures A2 and A3. We simulate 100 random subjects in our experimental design, choosing each option with 50% probability. Two patterns would be observed in our data if such random choice were prevalent. First, in the aggregate data, equalizing reductions would generally be low (the random first switch point would rarely stray above a few dollars) and would be insensitive to variation in probabilities or ranks. Second, in individual data, a wide degree of heterogeneity would be observed in the log difference, $\Delta\log(k)$, delivering apparent evidence of substantial rank dependence for many subjects.[10] These counterfactual predictions, along with the implication for the frequency of multiple switching, rule this hypothesis out as a plausible explanation of our data.

[8] For $\lambda > 2$, the CPE version of the KR model violates first-order stochastic dominance. As such, the case of $\lambda = 2$ represents the most extreme loss aversion possible without generating such behavior.

[9] Notably, these differences are in the opposite direction of those predicted by our calibrated CPT models. For our experimental values of q and p, $\pi(q) > \pi(p+q) - \pi(p)$ under standard parameterizations of the CPT model. Had we implemented our experiment with values of q and p for which this inequality was reversed at standard CPT parameterizations, the two models' predictions would have coincided.

[10] Where the simulated log difference exceeded the bounds of ± 3, we put the value at the boundary, including values of $\pm\infty$ when simulated as such.

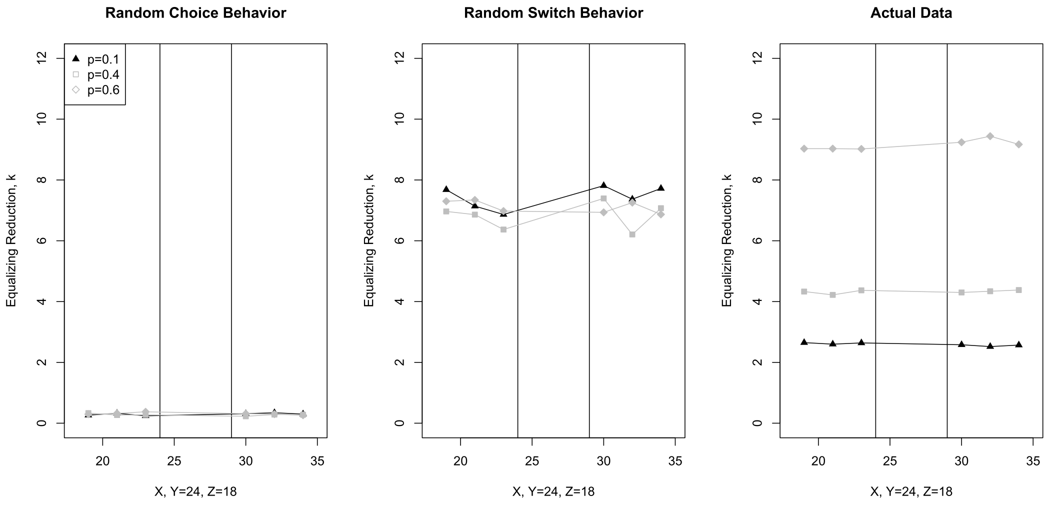

FIGURE A2.—Aggregate data with random response.

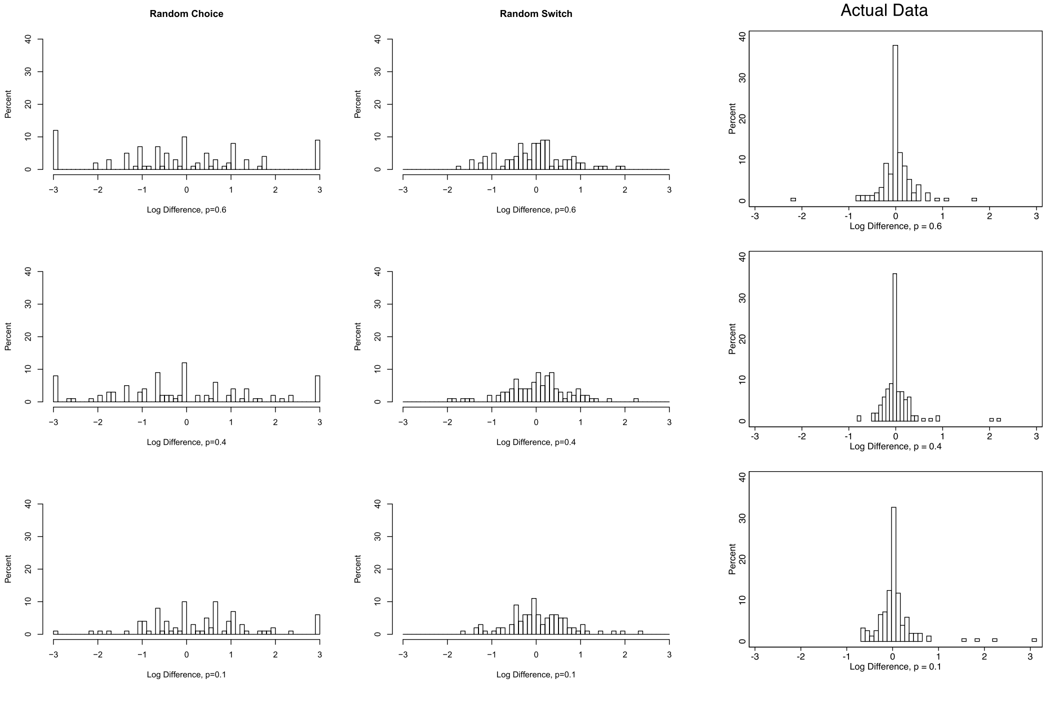

FIGURE A3.—Individual data with random response.

Second, we consider the possibility that each subject chooses a random switch point in each decision task. Simulated data for 100 such subjects appear in Figures A2 and A3. In addition to exhibiting no rank dependence, the aggregate choices of these subjects would be insensitive to probability distributions. At the individual level, we would again find wide heterogeneity in the log difference, $\Delta \log(k)$, providing apparent evidence of substantial rank dependence for many subjects. These implied patterns at the aggregate and individual levels clearly differ from the observed data. Our subjects respond to changes in probability distribution across tasks and exhibit subject-level log differences in equalizing reductions tightly centered around zero.

APPENDIX E: ADDITIONAL TABLES AND FIGURES

The following tables and figures are referenced in the main text and Supplemental Material Appendix F.

TABLE AII

CERTAINTY EQUIVALENTS[a]

	Certainty Equivalents (1)	Risk Premia (2)
$p = 0.05$	2.88 (0.19)	1.63 (0.19)
$p = 0.10$	3.83 (0.19)	1.33 (0.19)
$p = 0.25$	6.45 (0.17)	0.20 (0.17)
$p = 0.50$	10.72 (0.23)	−1.78 (0.23)
$p = 0.75$	15.44 (0.31)	−3.31 (0.31)
$p = 0.90$	19.83 (0.29)	−2.67 (0.29)
$p = 0.95$	21.63 (0.24)	−2.12 (0.24)

[a]Coefficients for certainty equivalents and risk premia calculated from interval regression of certainty equivalent on indicators for probability. Standard errors clustered on individual level in parentheses.

TABLE AIII

EQUALIZING REDUCTIONS WITHIN AND BETWEEN SUBJECTS[a]

	(1)	(2)	(3)	(4)	(5)	(6)
$\{p, q, 1-p-q\} =$ $\{0.4, 0.3, 0.3\}$	−4.72 (0.31)	−4.72 (0.17)	−5.03 (0.60)	−5.13 (0.60)	−5.13 (0.60)	−5.13 (0.60)
$\{p, q, 1-p-q\} =$ $\{0.1, 0.3, 0.6\}$	−6.40 (0.37)	−6.40 (0.18)	−6.65 (0.68)	−6.77 (0.68)	−6.77 (0.68)	−6.77 (0.68)
$(X > Y)$	0.26 (0.17)	0.26 (0.22)	−1.10 (0.85)	−0.93 (0.87)	−0.64 (0.83)	−0.74 (0.83)
$(X > Y) \times \{0.4, 3, 0.3\}$	−0.22 (0.16)	−0.22 (0.24)	0.73 (0.75)	0.60 (0.76)	0.60 (0.76)	0.60 (0.76)
$(X > Y) \times \{0.1, 3, 0.6\}$	−0.33 (0.18)	−0.33 (0.26)	0.82 (0.88)	0.71 (0.89)	0.71 (0.89)	0.71 (0.89)
$19 < \text{Age} < 22$					−0.10 (0.41)	−0.24 (0.43)
$\text{Age} \geq 22$					−0.33 (0.45)	−0.46 (0.46)
Male					0.89 (0.39)	0.98 (0.39)
Cognitve Reflect Test					0.41 (0.17)	0.40 (0.17)
Avg. Certainty Equivalent					0.19 (0.08)	
Constant	9.02 (0.39)	7.44 (0.59)	9.81 (0.65)	9.92 (0.65)	6.37 (1.17)	8.73 (0.77)
Predicted $\{0.6, 3, 0.1\}$	9.02 (0.39)	9.02 (0.16)	9.81 (0.65)	9.92 (0.65)	9.77 (0.60)	9.82 (0.60)
H_0: No Rank Dependence	$\chi^2(3) = 4.50$ $(p = 0.21)$	$\chi^2(3) = 1.82$ $(p = 0.61)$	$\chi^2(3) = 3.76$ $(p = 0.29)$	$\chi^2(3) = 2.50$ $(p = 0.47)$	$\chi^2(3) = 0.64$ $(p = 0.89)$	$\chi^2(3) = 0.86$ $(p = 0.84)$
Fixed Effects	No	Yes	No	No	No	No
First Block of Tasks Only	No	No	Yes	Yes	Yes	Yes
Demographic Controls	No	No	No	No	Yes	Yes
# Observations	2574	2574	429	405	405	405
# Subjects	143	143	143	135	135	135
Log-Likelihood	−8891.80	−8191.34	−1481.49	−1393.60	−1379.56	−1382.05

[a]Coefficients from interval regression of equalizing reduction on indicators for probability series $\{p, q, 1-p-q\}$ and order of outcome $X > Y$. Standard errors clustered at individual level in columns (1), (3), (4), (5), (6). Robust standard errors in parentheses in column (2). Column (4) restricts column (3) sample to 135 individuals with full control information. Constant, omitted category, is $\{p, q, 1-p-q\} = \{0.6, 3, 0.1\}$ with $X < Y$. Predicted average for $\{p, q, 1-p-q\} = \{0.6, 3, 0.1\}$ in (2), (5), (6) calculated as average of fixed effects or at the average level of controls. Tested null hypothesis of no rank dependence corresponds to test that coefficients $(X > Y)$, $(X > Y) \times \{0.4, 3, 0.3\}$, $(X > Y) \times \{0.1, 3, 0.6\}$ all equal zero.

TABLE AIV

EQUALIZING REDUCTIONS FOR ALL CONDITIONS[a]

	$\underline{k}$				$\overline{k}$			
$\{p, q, 1-p-q\}$	(1) $\underline{X}=19$	(2) $\underline{X}=21$	(3) $\underline{X}=23$	(4) $\underline{X}<Y$	(5) $\overline{X}=30$	(6) $\overline{X}=32$	(7) $\overline{X}=34$	(8) $\overline{X}>Y$
$\{0.6, 0.3, 0.1\}$	9.03	9.03	9.02	9.02	9.24	9.44	9.17	9.28
	(0.41)	(0.40)	(0.42)	(0.39)	(0.41)	(0.42)	(0.40)	(0.38)
$\{0.4, 0.3, 0.3\}$	4.33	4.22	4.37	4.31	4.30	4.34	4.38	4.34
	(0.14)	(0.13)	(0.14)	(0.12)	(0.14)	(0.15)	(0.13)	(0.12)
$\{0.1, 0.3, 0.6\}$	2.65	2.60	2.64	2.63	2.58	2.52	2.57	2.56
	(0.09)	(0.11)	(0.11)	(0.08)	(0.08)	(0.08)	(0.09)	(0.07)

[a] Coefficients calculated from interval regression of equalizing reduction on indicators for probability vector, value of $\underline{X}/\overline{X}$, and all interactions. Standard errors clustered on individual level in parentheses. Columns (4) and (8) provide estimated averages for $\underline{k}$ and $\overline{k}$ for columns (1)–(3) and (5)–(7), respectively.

TABLE AV

EQUALIZING REDUCTIONS WITH FIXED EFFECTS[a]

$\{p, q, 1-p-q\}$	$\underline{k}$	$\overline{k}$	$\Delta\widehat{\log(\frac{w_Y}{w_Z})}$ [95% Conf.]
$\{0.6, 0.3, 0.1\}$	9.02 (0.16)	9.28 (0.16)	0.03 (0.02) [−0.02, 0.08]
$\{0.4, 0.3, 0.3\}$	4.31 (0.07)	4.34 (0.07)	0.01 (0.02) [−0.04, 0.05]
$\{0.1, 0.3, 0.6\}$	2.63 (0.09)	2.56 (0.09)	−0.03 (0.05) [−0.12, 0.07]

[a] Mean behavior for $\underline{k}$ and $\overline{k}$ estimated from interval regression (Stewart (1983)) of experimental response on indicators for probability vector interacted with indicator for whether $X > Y$ with individual fixed effects. Constant taken as mean of fixed effects. Robust standard errors in parentheses.

TABLE AVI

EQUALIZING REDUCTIONS BETWEEN SUBJECTS ALTERNATE CONTROLS[a]

	Panel A: First Task Block (without Controls)			Panel B: First Task Block (with Alternate Controls)		
$\{p, q, 1-p-q\}$	$\underline{k}$	$\overline{k}$	$\Delta\widehat{\log(\frac{w_Y}{w_Z})}$ [95% Conf.]	$\underline{k}$	$\overline{k}$	$\Delta\widehat{\log(\frac{w_Y}{w_Z})}$ [95% Conf.]
$\{0.6, 0.3, 0.1\}$	9.81 (0.65)	8.71 (0.56)	−0.12 (0.09) [−0.30, 0.06]	9.82 (0.60)	9.09 (0.56)	−0.08 (0.09) [−0.25, 0.09]
$\{0.4, 0.3, 0.3\}$	4.78 (0.19)	4.41 (0.19)	−0.08 (0.06) [−0.20, 0.04]	4.70 (0.22)	4.56 (0.20)	−0.03 (0.07) [−0.16, 0.10]
$\{0.1, 0.3, 0.6\}$	3.16 (0.16)	2.88 (0.12)	−0.09 (0.07) [−0.22, 0.04]	3.06 (0.19)	3.03 (0.15)	−0.01 (0.08) [−0.17, 0.15]

[a] Mean behavior for $\underline{k}$ and $\overline{k}$ estimated from interval regression (Stewart (1983)) of experimental response on indicators for probability vector interacted with indicator for whether $X > Y$. Estimated change in relative decision weights, $\Delta\widehat{\log(w_Y/w_Z)}$, calculated as $\Delta\log(k)$. Standard errors clustered at individual level and calculated using the delta method, in parentheses. See Appendix Table AIII, columns (3) and (5) for detail. Panel A: No controls; 143 total subjects. Panel B: controls include age, gender, Cognitive Reflection Task score; 135 total subjects.

TABLE AVII

EQUALIZING REDUCTIONS WITH MULTIPLE SWITCHERS[a]

	$\underline{k}$				$\overline{k}$			
$\{p, q, 1-p-q\}$	(1) $\underline{X}=19$	(2) $\underline{X}=21$	(3) $\underline{X}=23$	(4) $\underline{X}<Y$	(5) $\overline{X}=30$	(6) $\overline{X}=32$	(7) $\overline{X}=34$	(8) $\overline{X}>Y$
{0.6, 0.3, 0.1}	8.72	8.78	8.69	8.73	8.92	9.09	8.76	8.93
	(0.41)	(0.38)	(0.41)	(0.38)	(0.40)	(0.42)	(0.40)	(0.38)
{0.4, 0.3, 0.3}	4.31	4.17	4.29	4.26	4.24	4.32	4.28	4.28
	(0.14)	(0.12)	(0.14)	(0.12)	(0.14)	(0.15)	(0.14)	(0.12)
{0.1, 0.3, 0.6}	2.62	2.56	2.58	2.59	2.59	2.55	2.59	2.58
	(0.09)	(0.11)	(0.11)	(0.08)	(0.09)	(0.08)	(0.09)	(0.07)

[a]Coefficients calculated from interval regression of equalizing reduction on indicators for probability vector, value of $\overline{X}/\underline{X}$, and all interactions. Standard errors clustered on individual level in parentheses. Columns (4) and (8) provide estimated averages for $\underline{k}$ and $\overline{k}$ for columns (1)–(3) and (5)–(7), respectively.

TABLE AVIII

EQUALIZING REDUCTIONS FIRST/LAST TASK BLOCK[a]

	$\underline{k}$				$\overline{k}$			
$\{p, q, 1-p-q\}$	(1) $\underline{X}=19$	(2) $\underline{X}=21$	(3) $\underline{X}=23$	(4) $\underline{X}<Y$	(5) $\overline{X}=30$	(6) $\overline{X}=32$	(7) $\overline{X}=34$	(8) $\overline{X}>Y$
Panel A: First Task Block								
{0.6, 0.3, 0.1}	11.10	8.01	10.49	9.81	7.87	9.39	8.85	8.71
	(1.14)	(0.99)	(1.13)	(0.65)	(1.12)	(1.21)	(0.72)	(0.56)
{0.4, 0.3, 0.3}	4.89	4.24	5.27	4.78	4.02	4.61	4.54	4.41
	(0.32)	(0.34)	(0.30)	(0.19)	(0.48)	(0.29)	(0.24)	(0.19)
{0.1, 0.3, 0.6}	3.17	3.08	3.24	3.16	2.62	2.89	3.03	2.88
	(0.25)	(0.25)	(0.35)	(0.16)	(0.20)	(0.11)	(0.20)	(0.12)
Panel B: Last Task Block								
{0.6, 0.3, 0.1}	9.46	11.37	6.85	9.12	8.09	8.53	9.72	8.75
	(0.77)	(0.93)	(0.95)	(0.54)	(0.93)	(1.11)	(1.24)	(0.64)
{0.4, 0.3, 0.3}	4.27	4.59	3.84	4.22	4.07	4.15	4.16	4.13
	(0.25)	(0.31)	(0.31)	(0.17)	(0.24)	(0.51)	(0.42)	(0.23)
{0.1, 0.3, 0.6}	2.63	2.37	2.60	2.55	2.51	2.56	2.37	2.48
	(0.19)	(0.23)	(0.28)	(0.13)	(0.14)	(0.34)	(0.17)	(0.14)

[a]Coefficients calculated from interval regression of equalizing reduction on indicators for probability vector, value of $\overline{X}/\underline{X}$, and all interactions. Standard errors clustered on individual level in parentheses. Columns (4) and (8) provide estimated averages for $\underline{k}$ and $\overline{k}$ for columns (1)–(3) and (5)–(7), respectively.

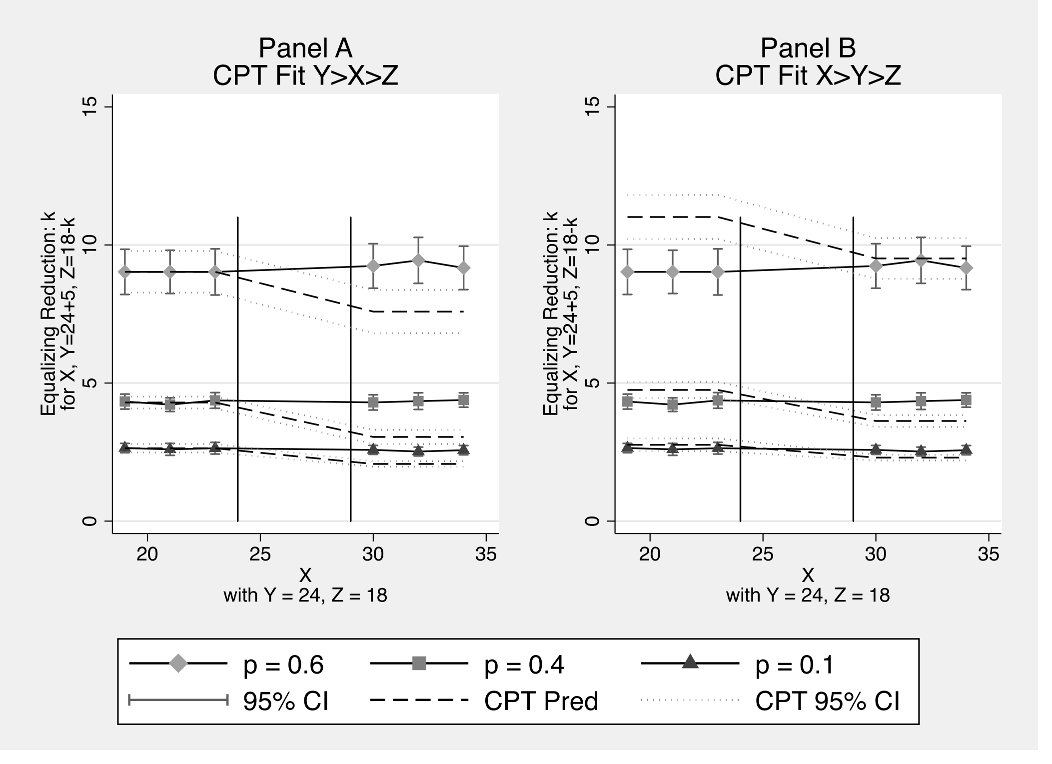

FIGURE A4.—Equalizing reductions with alternate benchmarks. *Notes*: Both panels: mean behavior for k estimated from interval regression of experimental response on indicators for probability vectors interacted with indicators for value of X. Standard errors clustered at individual level to provide 95% confidence interval. Supplemental Material Appendix Table AIV provides corresponding estimates. Dashed line corresponds to predicted values of equation (3) for CPT decisionmaker with risk preference parameters estimated from behavior. Panel A: predictions based on tasks with $Y > \underline{X} > Z$, $\alpha = 0.911$ (clustered s.e. $= 0.063$), and $\gamma = 0.784$ (0.020). Panel B: predictions based on tasks with $\overline{X} > Y > Z$, $\alpha = 1.024$ (0.082) and $\gamma = 0.830$ (0.022). Delta method used to provide 95% prediction confidence interval.

TASK 1

On this page you will make a series of decisions between two uncertain options. Option A will be a 40 in 100 chance of receiving $40, a 30 in 100 chance of receiving $36 and 30 in 100 chance of receiving $18. Initially Option B will be a 40 in 100 chance of receiving $44, a 30 in 100 chance of receiving $36 and 30 in 100 chance of receiving $18. As you proceed, Option B will change. For each row, decide whether you prefer Option A or Option B.

	Option A				*or*	Option B			
	40 in 100 Chance	30 in 100 Chance	30 in 100 Chance			40 in 100 Chance	30 in 100 Chance	30 in 100 Chance	
1)	$40	$36	$18	☐	*or*	$44	$36.00	$18.00	☐
2)	$40	$36	$18	☐	*or*	$44	$35.75	$17.75	☐
3)	$40	$36	$18	☐	*or*	$44	$35.50	$17.50	☐
4)	$40	$36	$18	☐	*or*	$44	$35.25	$17.25	☐
5)	$40	$36	$18	☐	*or*	$44	$35.00	$17.00	☐
6)	$40	$36	$18	☐	*or*	$44	$34.75	$16.75	☐
7)	$40	$36	$18	☐	*or*	$44	$34.50	$16.50	☐
8)	$40	$36	$18	☐	*or*	$44	$34.25	$16.25	☐
9)	$40	$36	$18	☐	*or*	$44	$34.00	$16.00	☐
10)	$40	$36	$18	☐	*or*	$44	$33.75	$15.75	☐
11)	$40	$36	$18	☐	*or*	$44	$33.50	$15.50	☐
12)	$40	$36	$18	☐	*or*	$44	$33.25	$15.25	☐
13)	$40	$36	$18	☐	*or*	$44	$33.00	$15.00	☐
14)	$40	$36	$18	☐	*or*	$44	$32.75	$14.75	☐
15)	$40	$36	$18	☐	*or*	$44	$32.50	$14.50	☐
16)	$40	$36	$18	☐	*or*	$44	$32.25	$14.25	☐
17)	$40	$36	$18	☐	*or*	$44	$32.00	$14.00	☐
18)	$40	$36	$18	☐	*or*	$44	$31.75	$13.75	☐
19)	$40	$36	$18	☐	*or*	$44	$31.50	$13.50	☐
20)	$40	$36	$18	☐	*or*	$44	$31.25	$13.25	☐
21)	$40	$36	$18	☐	*or*	$44	$31.00	$13.00	☐
22)	$40	$36	$18	☐	*or*	$44	$30.75	$12.75	☐
23)	$40	$36	$18	☐	*or*	$44	$30.50	$12.50	☐
24)	$40	$36	$18	☐	*or*	$44	$30.25	$12.25	☐
25)	$40	$36	$18	☐	*or*	$44	$30.00	$12.00	☐
26)	$40	$36	$18	☐	*or*	$44	$29.75	$11.75	☐
27)	$40	$36	$18	☐	*or*	$44	$29.50	$11.50	☐
28)	$40	$36	$18	☐	*or*	$44	$29.25	$11.25	☐
29)	$40	$36	$18	☐	*or*	$44	$29.00	$11.00	☐
30)	$40	$36	$18	☐	*or*	$44	$28.75	$10.75	☐
31)	$40	$36	$18	☐	*or*	$44	$28.50	$10.50	☐
32)	$40	$36	$18	☐	*or*	$44	$28.25	$10.25	☐
33)	$40	$36	$18	☐	*or*	$44	$28.00	$10.00	☐
34)	$40	$36	$18	☐	*or*	$44	$27.75	$9.75	☐
35)	$40	$36	$18	☐	*or*	$44	$27.50	$9.50	☐
36)	$40	$36	$18	☐	*or*	$44	$27.25	$9.25	☐
37)	$40	$36	$18	☐	*or*	$44	$27.00	$9.00	☐
38)	$40	$36	$18	☐	*or*	$44	$26.75	$8.75	☐
39)	$40	$36	$18	☐	*or*	$44	$26.50	$8.50	☐
40)	$40	$36	$18	☐	*or*	$44	$26.25	$8.25	☐

FIGURE A5.—Sample modified equalizing reduction.

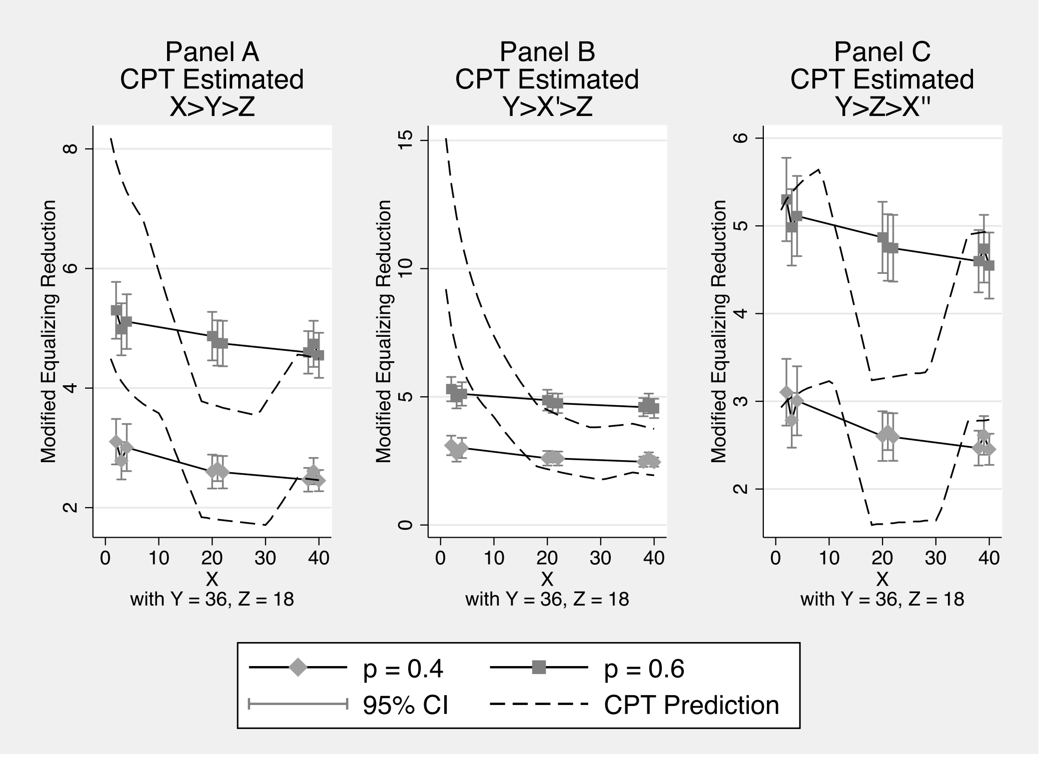

FIGURE A6.—Modified equalizing reductions with alternate benchmarks. *Notes*: All Panels: Mean behavior for modified equalizing reduction estimated from interval regression of experimental response on indicators for probability vectors interacted with indicators for value of X. Standard errors clustered at individual level to provide 95% confidence interval. Dashed line corresponds to predicted equalizing reductions for CPT decisionmaker with risk preference parameters estimated from behavior. Panel A: risk preferences estimated from tasks with $X > Y > Z$, $\alpha = 0.844$, and $\gamma = 0.785$. Panel B: risk preferences estimated from conditions with $Y > X' > Z$, $\alpha = 0.437$, and $\gamma = 0.863$. Panel C: risk preferences estimated from conditions with $Y > Z > X''$, $\alpha = 1.066$. and $\gamma = 0.741$.

APPENDIX F: ADDITIONAL ROBUSTNESS EXERCISES

F.1. *Alternative CPT Formulations*

Up to this point, we have focused exclusively on the Tversky and Kahneman (1992) parameterization of CPT. Others have proposed alternative functional forms. One leading alternative is due to Prelec (1998), who posited a probability weighting function of the form

$$\pi(p) = \exp\bigl(-\bigl(-\ln(p)\bigr)^{\gamma}\bigr).$$

To explore whether our conclusions are sensitive to functional form, we repeat our analysis for Prelec's specification. Using our data on certainty equivalents for binary lotteries, we arrive at the following estimates: weighting parameter $\gamma = 0.665$ (clustered s.e. $= 0.021$) and utility parameter $\alpha = 0.928$ (0.019). We then use the parameterized model to predict $\underline{k}$ and $\overline{k}$ as before. Results appear in Table AIX, Panel A. For convenience, we reproduce our results for Tversky and Kahneman's specification in Panel B.

TABLE AIX

EQUALIZING REDUCTION PREDICTIONS FOR ALTERNATIVE FUNCTIONAL FORMS[a]

	Panel A: Prelec Weighting				Panel B: Tversky Kahneman Weighting			
$\{p, q, 1-p-q\}$	$\underline{k}$	$\overline{k}$	$\Delta\log(k)$ [95% Conf.]	$\widehat{\Delta\log(\frac{w_Y}{w_Z})}$ [95% Conf.]	$\underline{k}$	$\overline{k}$	$\Delta\log(k)$ [95% Conf.]	$\widehat{\Delta\log(\frac{w_Y}{w_Z})}$ [95% Conf.]
{0.6, 0.3, 0.1}	8.77 (0.27)	6.50 (0.38)	−0.30 (0.03) [−0.35, −0.24]	−0.30 (0.03) [−0.36, −0.25]	7.58 (0.36)	6.06 (0.35)	−0.22 (0.01) [−0.25, −0.21]	−0.23 (0.01) [−0.24, −0.20]
{0.4, 0.3, 0.3}	4.81 (0.05)	2.63 (0.12)	−0.60 (0.04) [−0.69, −0.52]	−0.61 (0.04) [−0.70, −0.52]	4.01 (0.10)	2.52 (0.13)	−0.46 (0.03) [−0.52, −0.41]	−0.47 (0.03) [−0.52, −0.41]
{0.1, 0.3, 0.6}	2.95 (0.04)	1.70 (0.05)	−0.55 (0.04) [−0.64, −0.47]	−0.56 (0.04) [−0.64, −0.47]	2.65 (0.03)	1.87 (0.06)	−0.35 (0.03) [−0.40, −0.30]	−0.35 (0.03) [−0.40, −0.30]

[a]Panel A: Predicted behavior and change in decision weights calculated from equation (3) for Prelec CPT decisionmaker with parameters $\alpha = 0.928$ (s.e. $= 0.019$) and $\gamma = 0.665$ (0.021). Standard errors clustered at individual level and calculated using the delta method, in parentheses. Panel B: Predicted behavior and change in decision weights calculated from equation (3) for Kahneman and Tversky CPT decisionmaker with parameters $\alpha = 0.965$ (s.e. $= 0.021$) and $\gamma = 0.703$ (0.015). Standard errors clustered at individual level and calculated using the delta method, in parentheses.

Note that the predicted discontinuities are even larger, and hence less consistent with actual behavior, with the Prelec specification.

F.2. *Using Explicit Rank Changes*

The last task block in each session featured $X = \$25$ and $Y = \$24$, so that adding $m = \$5$ to Y changes its rank. Using the estimated aggregate CPT parameter values, one predicts equalizing reductions of 7.28, 3.71, and 2.49 for $\{p, q, 1-p-q\} = \{0.6, 0.3, 0.1\}$, $\{0.4, 0.3, 0.3\}$, and $\{0.1, 0.3, 0.6\}$, respectively. Note that these values are close to the CPT predictions of $\underline{k}$ reported in Table III, Panel B and are substantially higher than those of $\overline{k}$.

For $\{p, q, 1-p-q\} = \{0.6, 0.3, 0.1\}$, the mean equalizing reduction is 8.94 (clustered s.e. $= 0.41$). This value is statistically indistinguishable from the actual value of $\underline{k}$ for $X' < Y$ reported in Table III, Panel A, $\chi^2(1) = 0.27$ ($p = 0.61$), and is significantly lower than the value of $\overline{k}$ for $X > Y$, $\chi^2(1) = 3.44$ ($p = 0.06$). For $\{p, q, 1-p-q\} = \{0.4, 0.3, 0.3\}$, the mean equalizing reduction is 4.12 (0.13), significantly lower than the values of both $\underline{k}$ and $\overline{k}$ reported in Table III, Panel A, $\chi^2(1) = 4.19$ ($p = 0.04$) and $\chi^2(1) = 5.36$ ($p = 0.02$), respectively. For $\{p, q, 1-p-q\} = \{0.1, 0.3, 0.6\}$, the mean equalizing reduction is 2.34 (0.08), significantly lower than the values of both $\underline{k}$ and $\overline{k}$ reported in Table III, Panel A, $\chi^2(1) = 18.82$ ($p < 0.01$) and $\chi^2(1) = 11.55$ ($p < 0.01$), respectively.

The pattern described in the previous paragraph is, on its face, somewhat puzzling. If the equalizing reduction does not depend on the ranking of the payoff Y, it is difficult to see why it should be systematically lower in the transitional region. Certainly, that implication is inconsistent not only with CPT, but also with PT and EU. A possible explanation is that the $X = 25$ task block always comes last, and equalizing reductions decline as the experiment progresses from the first task block to the last (see Table AVIII, Panel B). Consistent with this hypothesis, the equalizing reductions in the $X = \$25$ tasks are quite close to the values reported for those for the last task block (see Table AVIII, Panel B).

F.3. *Multiple Switching*

Our main results are derived from the choices of 143 subjects who did not exhibit multiple switching in any task. For Table AX, we include the remaining subjects, each of whom exhibited multiple switching at least once. The results are qualitatively unchanged. As in Table III, we predict substantial differences between $\underline{k}$ and $\overline{k}$ but observe none.[11] Thus, our conclusions are robust with respect to the inclusion or exclusion of potentially confused subjects.

[11]Supplemental Material Appendix Table AVII provides estimates of equalizing reductions for each value of X and X', and demonstrates the stability of equalizing reductions across these values.

TABLE AX

EQUALIZING REDUCTIONS WITH MULTIPLE SWITCHERS[a]

	Panel A: Mean Behavior and Estimated Rank Dependence			Panel B: CPT Estimates and Predicted Rank Dependence		
				Equalizing Reductions $Y > \underline{X} > Z$ $\gamma = 0.776\ (0.020)$	Equalizing Reductions $\overline{X} > Y > Z$ $\gamma = 0.813\ (0.023)$	Certainty Equivalents $\gamma = 0.703\ (0.015)$
$\{p, q, 1-p-q\}$	$\underline{k}$	$\overline{k}$	$\Delta \widehat{\log(\frac{w_Y}{w_Z})}$ [95% Conf.]	$\Delta \widehat{\log(\frac{w_Y}{w_Z})}$ [95% Conf.]	$\Delta \widehat{\log(\frac{w_Y}{w_Z})}$ [95% Conf.]	$\Delta \widehat{\log(\frac{w_Y}{w_Z})}$ [95% Conf.]
{0.6, 0.3, 0.1}	8.73 (0.38)	8.93 (0.38)	0.02 (0.02) [−0.01, 0.06]	−0.19 (0.01) [−0.21, −0.16]	−0.16 (0.02) [−0.19, −0.12]	−0.23 (0.01) [−0.25, −0.22]
{0.4, 0.3, 0.3}	4.26 (0.12)	4.28 (0.12)	0.01 (0.02) [−0.03, 0.04]	−0.36 (0.03) [−0.43, −0.29]	−0.30 (0.04) [−0.37, −0.22]	−0.49 (0.03) [−0.54, −0.44]
{0.1, 0.3, 0.6}	2.59 (0.08)	2.58 (0.07)	−0.00 (0.02) [−0.05, 0.04]	−0.26 (0.03) [−0.31, −0.20]	−0.21 (0.03) [−0.26, −0.15]	−0.37 (0.03) [−0.42, −0.32]

[a]Panel A: Mean behavior for $\underline{k}$ and $\overline{k}$ estimated from interval regression (Stewart (1983)) of experimental response on indicators for probability vector interacted with indicator for whether $X > Y$. Estimated change in relative decision weights, $\Delta \widehat{\log(w_Y/w_Z)}$, calculated as $\Delta \log(k)$. Standard errors clustered at individual level and calculated using the delta method, in parentheses. See Supplemental Material Appendix Table AIII, column (1) and Supplemental Material Appendix Table AIV for detail. Panel B: Predicted change in probability weight for CPT decisionmaker with probability weighting estimated solely from equalizing reductions with $Y > \underline{X} > Z$, from equalizing reductions with $\overline{X} > Y > Z$, or from certainty equivalents data. Estimated probability weighting parameter noted for each prediction. Estimated change in relative decision weights, $\Delta \widehat{\log(w_Y/w_Z)}$, calculated as $\log(\pi(p+q) - \pi(p)) - \log(\pi(q))$ for estimated weighting function. Standard errors clustered at individual level and calculated using the delta method, in parentheses.

REFERENCES

BELL, D. E. (1985): "Disappointment in Decision Making Under Uncertainty," *Operations Research*, 33 (1), 1–27. [7]

GONZALEZ, R., AND G. WU (1999): "On the Shape of the Probability Weighting Function," *Cognitive Psychology*, 38, 129–166. [1]

KAHNEMAN, D., AND A. TVERSKY (1979): "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47 (2), 263–291. [1]

KOSZEGI, B., AND M. RABIN (2006): "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics*, 121 (4), 1133–1165. [9]

——— (2007): "Reference-Dependent Risk Attitudes," *The American Economic Review*, 97 (4), 1047–1073. [9]

LOOMES, G., AND R. SUGDEN (1986): "Disappointment and Dynamic Consistency in Choice Under Uncertainty," *Review of Economic Studies*, 53 (2), 271–282. [7]

PRELEC, D. (1998): "The Probability Weighting Function," *Econometrica*, 66 (3), 497–527. [19]

STEWART, M. B. (1983): "On Least Squares Estimation When the Dependent Variable Is Grouped," *The Review of Economic Studies*, 50 (4), 737–753. [15,22]

TVERSKY, A., AND C. R. FOX (1995): "Weighing Risk and Uncertainty," *Psychological Review*, 102 (2), 269–283. [1]

TVERSKY, A., AND D. KAHNEMAN (1992): "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5 (4), 297–323. [1,2,5,19]

WAKKER, P., I. EREV, AND E. U. WEBER (1994): "Comonotonic Independence: The Critical Test Between Classical and Rank Dependent Utility Theories," *Journal of Risk and Uncertainty*, 9, 195–230. [1,3]

WU, G. (1994): "An Empirical Test of Ordinal Independence," *Journal of Risk and Uncertainty*, 9, 39–60. [1, 3]

Co-editor Itzhak Gilboa handled this manuscript.

Manuscript received 3 September, 2018; final version accepted 11 January, 2020; available online 22 January, 2020.