Is There a Plausible Theory for Decision under Risk? A Dual Calibration Critique

By James C. Cox, Vjollca Sadiraj, Bodo Vogt, and Utteeyo Dasgupta*

Experimental Economics Center

Georgia State University

June 2010

* Financial support was provided by the National Science Foundation (grant numbers IIS-0630805 and SES-0849590).

Is There a Plausible Theory for Decision under Risk? A Dual Calibration Critique

Abstract: Theories of decision under risk that assume decreasing marginal utility of money have been critiqued with concavity calibration arguments. Since that critique uses varying payoffs and fixed probabilities, it cannot have implications for calibration of nonlinear probability transformation, which is another way to model risk aversion. The concavity calibration critique also has no implication for theories with variable reference points. This paper introduces a new type of (varying-probabilities, fixed-payoffs) calibration that applies to nonlinear transformation of probabilities. It also applies to theories with constant or variable reference points. The two types of calibrations yield dual paradoxes: a pattern of risk aversion that conforms to the (resp. dual) independence axiom implies implausible risk aversion for theories with functionals that are linear in payoffs (resp. probabilities). Functionals that are nonlinear in both payoffs and probabilities are subject to both types of calibration critique. The dual calibrations make clear why plausibility problems with theories of decision under risk may be fundamental. They are fundamental if their empirical relevance can be demonstrated. This paper reports seven experiments that provide data on the empirical relevance of the dual calibration critique of decision theory. (JEL C90, D81)

Can prominent theories of decision under risk rationalize both small-stakes risk aversion and large-stakes risk aversion? How do loss aversion and reference payoffs enter in the answer to this question? Can some existing theories, but not others, rationalize same-stakes (i.e. small-stakes *or* large-stakes) risk aversion? We offer a theoretical duality approach that addresses these questions. We present two (dual) paradoxes in which patterns of risk aversion that conform to one theory of decision under risk imply implausible risk aversion in the dual to that theory. One wonders then whether data conform to either or both of the dual "calibration patterns" for which prominent theories imply implausible risk aversion. We report seven experiments that address that question.

Rabin (2000) sparked the literature on concavity calibration by identifying a varying-payoffs pattern of small-stakes risk aversion that, through calibration arguments, can be shown to imply implausible large-stakes risk aversion for the expected utility of terminal wealth model. Several subsequent authors extended Rabin's varying-payoffs, concavity calibration analysis to apply to a class of theories that assume decreasing marginal utility of money. How fundamental is this challenge to the plausibility of theories of decision under risk? We address this question about fundamentality both theoretically and empirically.

Our theoretical discussion is based on duality. We explain in this paper that the varying-payoffs patterns of small-stakes risk aversion used in calibrations by Rabin (2000) and all subsequent authors conform to the *dual* independence axiom (Yaari, 1987). This presents a paradox: patterns of risk aversion that *characterize* rational behavior for the dual theory of expected utility (Yaari, 1987), with constant marginal utility of money for all risk preferences, imply implausible risk aversion for theories with decreasing marginal utility of money such as expected utility theory and rank dependent utility theories (Quiggin, 1993, Tversky and Kahneman, 1992). One wonders then whether there are varying-probabilities patterns of risk aversion that conform to the independence axiom of expected utility theory and yet imply implausible risk aversion for theories with nonlinear transformation of probabilities such as the dual theory of expected utility and cumulative prospect theory. We explain that the answer to this question is "yes." This presents a second (dual) paradox: patterns of risk aversion that *characterize* rational behavior for expected utility theory, with linearity in probabilities for all risk preferences, imply implausible risk aversion for theories that represent risk averse preferences with nonlinear probability transformations (with or without nonlinear transformation of payoffs).

Previous literature on concavity calibration has focused on the inability of some prominent theories to rationalize *both* small-stakes risk preferences and large-stakes risk preferences. This leaves open the question of whether calibration arguments have implications for decision theories' ability to rationalize same-stakes (i.e. small-stakes *or* large-stakes) risk preferences. In other words, if a researcher is content to view existing applications of decision theories as having only same-stakes implications, then does (s)he escape criticism based on calibration arguments? We explain that theories that represent risk aversion with nonlinear probability transformations have implausible implications even for same-stakes risk preferences.

Of course the fundamentality of the above theoretical results rests on empirical validity of the patterns of risk aversion used in the calibration propositions. To date, however, there has been only argument about the "reasonableness" of the calibration suppositions but no data from real-payoff, controlled experiments to inform the issue. We explain why researchers encounter especially difficult problems in conducting experiments to test the empirical validity of suppositions in calibration propositions and discuss solutions to these problems that were implemented in our experiments. The paper reports seven experiments conducted over several years in three countries (India, Germany, and the United States) with idiosyncratic opportunities for implementing a variety of experimental designs and protocols covering both varying-payoffs and varying-probabilities calibration patterns of risk aversion that have implications for theories of decision under risk.

Previous literature reports varying-payoffs calibration patterns that apply to models defined on (a) terminal wealth or (b) income. Studies that focus on terminal wealth models include Rabin (2000), Neilson (2001), and Safra and Segal (2008). Rabin demonstrated results that apply to the expected utility of terminal wealth model. Neilson showed that Rabin's concavity calibration critique applies to rank-dependent utility of terminal wealth. Safra and Segal introduced a stochastic version of Rabin's calibration pattern that produces anomalies for additional non-expected utility models in which preferences are defined on terminal wealth.

None of the above calibration propositions apply to theories that incorporate loss aversion because the reference points for terminal wealth models are amounts of initial wealth not amounts of income (that define losses and gains). Calibrations for models defined on income are reported by Barberis, Huang, and Thaler (2006), Cox and Sadiraj (2006), and Rubinstein (2006). Barberis, Huang, and Thaler examined the implications of calibration for recursive utility with first-order and second-order risk aversion. Cox and Sadiraj looked at calibration issues for Tversky and Kahneman's (1992) cumulative prospect theory and two expected utility models that are alternatives to the terminal wealth model. Rubinstein took the concavity calibration critique to time preferences under risk.

All of the previous studies were built on the same varying-payoffs pattern of small-stakes risk aversion that first appeared in Rabin (2000). As we explain, the varying-payoffs calibration patterns in previous literature have no implausible risk aversion implications for the dual theory of expected utility (Yaari, 1987), an early alternative to expected utility theory that models risk aversion (solely) with nonlinear transformation of probabilities. As explained by Wakker (2005), those calibration patterns have no implausible risk aversion implications for recent versions of cumulative prospect theory with variable reference amounts of income.

In this paper, we introduce a varying-probabilities pattern of risk aversion that has calibration implications for theories that incorporate risk aversion with nonlinear transformation of probabilities (with or without nonlinear transformation of payoffs). The new calibration does have implausible risk aversion implications for the dual theory of expected utility and for cumulative prospect theory with variable reference amounts of income and loss aversion. The new calibration pattern implies *same-stakes* (as well as large-stakes vs. small-stakes) implausible risk aversion for theories that incorporate nonlinear transformation of probabilities. As a result, such theories are called into question even for applications that preserve the domain of payoffs.

We report dual calibration propositions and several corollaries. Proposition 1 identifies varying-probabilities patterns of risk aversion that have implausible (small-stakes vs. large-stakes *and* same-stakes) risk aversion implications for dual theory of expected utility. Proposition 2 uses varying-payoffs patterns, appearing in previous literature, that have implausible small-stakes vs. large-stakes risk aversion implications for expected utility theory. Each proposition has a corollary that extends the calibration to rank-dependent theories, including cumulative prospect theory, that model risk preferences with nonlinear transformations of both probabilities and payoffs. In this way, such theories are shown to be subject to both types of calibration critique. The new, varying-probabilities calibration also applies to theories with nonlinear transformations of probabilities and variable reference amounts of income (with or without loss aversion).

I. Independence, Dual Independence, and Calibration Patterns

We start with two examples that illustrate dual calibration paradoxes. The first example, known as Rabin's pattern, is a pair of risk preference statements that can be rationalized by the dual theory of expected utility (DTEU) but cannot be rationalized by expected utility theory (EUT). The second example introduces a new pair of risk preference statements that can be rationalized by EUT but cannot be rationalized by DTEU. These examples illustrate patterns in the dual calibration propositions reported in sections II and III. Both patterns have implications for theories that incorporate nonlinear transformations of both payoffs and probabilities.

A. An Example of Calibration for Varying Payoffs

Consider a representative example from previous literature (Rabin, 2000) consisting of Statement P.2 (a pattern of small-stakes risk aversion) and Statement Q.2 (a large-stakes lottery preference). Statement P.2 says that the agent rejects the 50/50 lottery with loss of 100 or gain of 105 at all amounts of initial wealth w between 100 and 300,000.[1] Statement Q.2 says that the agent prefers the 50/50 lottery that pays 0 or 5 million to getting 10,000 for sure at initial wealth 290,000. Rabin shows that Statement P.2 is inconsistent with Statement Q.2. So, the expected

[1] *Sections III and V explore the implications of varying the size of the payoff interval over which P.2 holds.*

utility of terminal wealth model cannot rationalize *both* of these statements about risk preferences; that is, this model is inconsistent either with Statement P.2 or with Statement Q.2.

In contrast, DTEU can rationalize risk preferences that satisfy both Statements P.2 and Q.2. The Statement P.2 pattern of small-stakes risk aversion conforms to the dual independence axiom: according to this axiom, a DTEU agent who rejects the 50/50 lottery with payoffs of -100 or $+105$ for *some* value of initial wealth w must reject the same lottery for *all* values of w. An easy way to see this implication is through the linearity in payoffs property that characterizes the DTEU functional (as a consequence of the dual independence axiom).

Paradoxically, the pattern of small stakes risk aversion contained in Statement P.2: (a) implies *implausible* large-stakes risk aversion (negation of statement Q.2) for EUT; but (b) conforms to rational behavior for DTEU because it conforms to the dual independence axiom. It has *no* implication of implausible large stakes risk aversion for DTEU.

B. An Example of Calibration for Varying Probabilities

Here we introduce a pair of risk preference statements that cannot be rationalized by DTEU but can be rationalized by EUT. Consider an agent with some initial wealth $\overline{w}$ between 0 and 300,000 who (weakly) prefers 1 million for sure to a 50/50 lottery that pays 2.5 million or 0. Then it is a straightforward implication of linearity in probabilities of the EUT functional that EUT implies that this agent prefers a three outcome lottery that pays 2.5 million or 1 million or 0, with probabilities $p-0.05$ and 0.1 and $1-p-0.05$, to a two outcome lottery that pays 2.5 million or 0, with probabilities p and $1-p$, for all $p \in \{0.05, 0.1, ..., 0.9, 0.95\}$. Although such risk preferences conform to the independence axiom of EUT they have implausible risk aversion implications for DTEU, as we shall now explain.

Let Statement P.1 say that the agent rejects a lottery that pays 2.5 million or 0 with probabilities p and $1-p$ in favor of a lottery that pays 2.5 million or 1 million or 0 with probabilities $p-0.05$ and 0.1 and $1-p-0.05$ for all $p \in \{.05, 0.1, ..., 0.90, 0.95\}$.[2] Statement Q.1 says that the agent prefers the 50/50 lottery that pays 0 or 58,665 to getting 1,000 for sure. Proposition 1, below, shows that DTEU is inconsistent either with Statement P.1 or Statement Q.1. The implausibility of the risk preferences in Statement Q.1 is increasing with the number of

---

[2] *Sections II and V explore the implications of varying the size of the probability interval over which P.1 holds.*

probability sub-intervals in Statement P.1. For example, if Statement P.1b is that an agent prefers the three outcome lottery to the two outcome lottery for all $p \in \{.01, 0.02, ..., 0.98, 0.99\}$ then, according to DTEU, the agent will prefer getting 1,000 for sure to a 50/50 lottery that pays 0 or 633 billion. The Statement Q.1b that is inconsistent with Statement P.1b according to DTEU is: the agent prefers the 50/50 lottery that pays 0 or 633 billion to getting 1,000 for sure. In this way, DTEU cannot rationalize such risk preferences.

Paradoxically, the Statement P.1 pattern of small-stakes risk aversion: (a) implies *implausible* large-stakes risk aversion (the negation of statement Q.1) for DTEU; but (b) conforms to rational behavior for EUT because it conforms to the independence axiom.

Theories such as cumulative prospect theory with functionals that exhibit both nonlinearity in payoffs and nonlinearity in probabilities are inconsistent with slightly modified versions of *both* pairs of statements, (P.1,Q.1) and (P.2,Q.2), as explained in sections II and III.

II. Calibrations with Varying Probabilities (and Fixed Payoffs)

We introduce a calibration proposition for the dual theory of expected utility and corollaries that apply to theories with functionals that are nonlinear in both probabilities and payoffs. The design of experiments reported in section VI is based on calibration patterns discussed here.

A. Calibration for Linear Money Transformation Functions

Let $\{y_m, p_m; ...; y_2, p_2; y_1\}$, denote an *m*-outcome lottery that pays y_k with probability p_k, for $k = 2, \cdots, m$, and pays y_1 with probability $1 - \sum_{k=2}^{m} p_k$. We use the convention $y_{k+1} \geq y_k$, for $k = 1, \cdots, m-1$. Whenever the smallest payoff is zero (i.e., $y_1 = 0$), we use the simpler notation $\{y_m, p_m; ...; y_2, p_2\}$.

Consider the $2n-1$ pairs of lotteries $A_i = \{y, p_i\}$, and $B_i = \{y, p_i - \delta; x, 2\delta\}$, where $p_i = i/2n$, $\delta = 1/2n$, and $i = 1, 2, \cdots, 2n-1$. In each pair of lotteries, lottery B_i is constructed from lottery A_i by transferring probability mass $1/2n$ from both the highest payoff y and the lowest payoff 0 to the intermediate payoff x.

Suppose that an agent prefers the three outcome lottery B_i to the two outcome lottery A_i, for all $i = 1, 2, \cdots, 2n-1$. Note that, by the independence axiom, any expected utility maximizer who prefers x for sure to the 50/50 lottery that pays y or 0 satisfies this supposition. Proposition 1 shows that if the high outcome y is larger than twice the intermediate outcome x then this supposition implies implausible risk aversion for DTEU agents. The following standard notation is used: $\succsim$ indicates weak preference; $\succ$ indicates strong preference; and N denotes the set of positive integers. Define $K(t,m,n) = 1 + \sum_{j=1}^{m} (t-1)^{j} \Big/ \sum_{i=1}^{n} (t-1)^{1-i}$.

Proposition 1. *Let* $n \in N$ and $y > 2x > 0$ be given. Let

$p_i = i/2n, \delta = 1/2n$ and $G = K(y/x, n, n)$. Consider the statements

P.1 $\{y, p_i - \delta; x, 2\delta\} \succsim \{y, p_i\}$, for all $i \in \{1, 2, \cdots, 2n-1\}$ and

Q.1 $\{zG, 0.5\} \succ z$, for some $z > 0$.

a. Any EUT agent who prefers x to $\{y, 0.5\}$ satisfies P.1.
b. There are EUT agents who satisfy both P.1 and Q.1.
c. There are no DTEU agents who satisfy both P.1 and Q.1.

Proof: see appendix A.2.

Note that $G = K(y/x, n, n) \underset{n->\infty}{\rightarrow} \infty$ for $y/x > 2$. Hence, the larger the value of n, the more extreme the implications of the P.1 pattern of risk aversion. Put differently, for any G, as big as one chooses, there exists n such that for weak preference for the three outcome lottery B_i over the two outcome lottery A_i, for all $i \in \{1, 2, \cdots, 2n-1\}$, DTEU predicts a preference for z for sure over the risky lottery $\{zG, 0.5\}$ for all $z > 0$.[3]

Some numerical illustrations of Proposition 1 are reported in Table 1. In the table, C = y/x, the ratio of the highest payoff to the second highest payoff in the three prize lottery. With C = 2.5 and n = 20 Proposition 1 tells us that for this P.1 pattern DTEU predicts that the agent prefers 1,000 for sure to a lottery that pays 3.3 million or 0 with even odds, as reported in the first column and third row of Table 1. With C = 3.5 and n = 50 the prediction is preference for 1,000 for sure over a 50/50 lottery that pays 0 or 0.78×10^{23}. Finally, with $C = 5$ and n = 10, the prediction is preference for 1,000 for sure to the 50/50 lottery that pays 0 or 1 billion.

B. Calibration for Linear and Nonlinear Money Transformation Functions

Proposition 1 is stated for the dual theory of expected utility that is characterized by a preference functional that is linear *in payoffs* for all risk preferences. The generalization is straightforward for a class of models with nonlinear transformation of decumulative probabilities (referred to as NTDP) that represent risk preferences with linear or nonlinear transformation of payoffs $\upsilon(\cdot)$ as well as nonlinear transformation of probabilities. First consider NTDP with constant, zero-income reference point, as in Tversky and Kahneman (1992). For $\upsilon(\cdot)$ sub-additive on positive payoffs one has:

Corollary 1.1. For $\upsilon(y) > 2\upsilon(x)$, there are no NTDP agents with zero-income reference point who satisfy both P.1 and Q.1 with $G_\upsilon = K(\upsilon(y)/\upsilon(x), n, n)$.

Proof: see appendix A.2.

It can be verified that for $\upsilon(y)/\upsilon(x) > 2$, $\lim_{n\to\infty} G_\upsilon = \lim_{n\to\infty} K(\upsilon(y)/\upsilon(x), n, n) = \infty$. Implications of Corollary 1.1 are given in Table 1 for the (alternative) definition $C = \upsilon(y)/\upsilon(x)$. For example, if the high payoff y is *k* times as large as the intermediate (positive) payoff x and the concave value function of (positive) payoffs is such that $\upsilon(kx)/\upsilon(x) \geq 3$ then implications of calibration pattern P.1 are given by the $C = 3$ column of Table 1, and so on.

A reference-dependent theory such as prospect theory can incorporate variable reference amounts of money payoff. Wakker (2005) argues that variable reference points can immunize prospect theory to concavity calibration arguments based on the small-stakes risk aversion pattern introduced by Rabin (2000). In contrast, the dual calibration pattern introduced herein is robust to variable reference amounts of income. The reason for this is straightforward: the calibration is constructed by varying the probabilities for which three or two payoffs are paid, not by varying the payoff amounts. Hence it makes no difference to the calibration reported here whether the reference amount of payoff is or is not fixed at zero payoff. Here is a formal statement of the result. Let $\mu(\cdot) < 0$ denote the value function for negative payoffs and define $R = -\upsilon(y-x)/\mu(-x)$. For $\upsilon(\cdot)$ sub-additive on positive payoffs one has:

---

[3] Note that this proposition makes no explicit assumption on the curvature of the probability transformation.

Corollary 1.2. Let the reference point be the intermediate payoff x and $\upsilon(y-x) > -\mu(-x)$. There are no loss averse NTDP agents who satisfy both P.1 and Q.1 with $G_R = K(R+1,n,n)$.
Proof: see appendix A.2.

Note that for $R > 1$, $G_R = K(R+1,n,n) \to \infty$ as $n \to \infty$. Corollary 1.2 holds for both nonlinear and piece-wise linear value functions. Similar corollaries can be stated for cases in which the reference point is the highest or lowest payoff rather than the intermediate payoff.

C. Calibration for Proper Subsets of Discrete Probabilities in [0,1]

Proposition 1, Corollary 1.1, and Corollary 1.2 report the implications of preference for the three-outcome lottery over the two-outcome lottery *for all* probabilities, $p_i = i/2n \in [0,1]$ of the high payoff from 0 to $1-1/2n$. But what if statement P.1 is not true *for all* $p_i \in [0,1]$? Perhaps the preference for the three outcome lottery over the two outcome lottery holds only for a proper subset of the [0,1] interval of probabilities. This question is addressed by Corollary 1.3.

Some examples of questions addressed by Corollary 1.3 are the following. What if statement P.1 is not true for all p but only for all $p \in \{0.50, 0.51, ..., 0.98, 0.99\}$? (This could occur for some patterns of non-EUT indifference curves in the Marschak-Machina triangle diagram that fan out.) Then, according to Corollary 1.3, DTEU predicts that the agent prefers a 50/50 lottery that pays 1 or 25,253 to a 25/75 lottery that pays 0 or 25,253, which is clearly implausible risk aversion. So statement Q.1 in this case is preference for the 25/75 lottery with outcomes 0 and 25,253 to the 50/50 lottery with outcomes 1 and 25,253. Another example involves the case of statement P.1 being satisfied only for all $p \in \{0.01, 0.02..., 0.49, 0.50\}$. (This could occur for some patterns of non-EUT indifference curves in the Marschak-Machina triangle diagram that fan in.) Then, according to Corollary 1.3, DTEU predicts that the agent prefers a 50/50 lottery that pays 0 or 1,000 to a 75/25 lottery that pays 0 or 25 million, which is clearly implausible risk aversion. Therefore, statement Q.1 in this last example is preference for the 75/25 lottery that pays 0 or 25 million over the 50/50 lottery with outcomes 0 and 1,000.

Consider the class of models NTDP for which C denotes: y/x for a functional that is linear in payoffs or $(y-x)/x$ for a functional that is piecewise linear in payoffs, with discontinuous slope (loss aversion) at x, or $\upsilon(y)/\upsilon(x)$ for a functional that is nonlinear in payoffs or $-\upsilon(y-x)/\mu(-x)+1$ for a functional that is nonlinear in payoffs with discontinuous slope at x. Without any loss of generality let $\upsilon(0)=0$. For $\upsilon(\cdot)$ sub-additive on positive payoffs one has:

Corollary 1.3. Denote $G=K(C,n/2,n/2)$, for an even n, and $G'=K(C,m,n)$ such that $\{m\in N$ and $m<n\}$. There are no NTDP agents who satisfy both:

a. P.1 for all $i\in\{1,\cdots,n,...,n+m\}$ *and* Q.1 with G'

b. P.1 for all $i\in\{n,...,2n-1\}$ *and* Q*.1: $\{zG,0.75\}\succ\{zG,0.5;z\}$, for some $z>0$.

c. P.1 for all $i\in\{1,\cdots,n\}$ *and* Q**.1: $\{zG,0.25\}\succ\{z,0.5\}$, for some $z>0$.

Proof: see Appendix A.2

Part *a* of the corollary states implications for the case when the interval of preference for the three outcome lottery over the two outcome lottery is a subset of (0,1) that includes (0,1/5]. Part *b* states results for the case where the interval of preference is [0.5,1). Finally, part *c* states results for the case when the interval of preference for the three outcome lottery is (0,0.5]. In analysis of data from the experiments, we will also need to know the implications of preference for the three-outcome lottery over the two-outcome lottery *for only some* of the probabilities in an experiment design. Corollary 1.3 will be applied in analysis of data in section VI.

D. Implausible Same-Stakes Risk Aversion

Implausible risk aversion implications of theories that transform probabilities are not limited to large-stakes vs. small-stakes comparisons. Such theories also predict implausible same-stakes risk aversion. There can be somewhat different ideas of what might be meant by "same-stakes." One natural definition is that the payoffs in the lotteries in statement Q.1 are weakly between the highest and lowest payoffs in the lotteries in statement P.1, that is, they are in the same payoff domain of application of the theory. Proposition 1 and its corollaries imply such "same-stakes" implausible risk aversion, as can be seen from the following.

Statement P.1 involves lotteries with high payoff amount y, intermediate payoff amount x, and low payoff amount 0. Calibration implications are derived for different ratios of $y/x=C$. Statement Q.1 says that for a sufficiently large G, a 50/50 lottery that pays zG or 0 is preferred to z. As explained in section II.*a* and Table 1, the value of G can be set as large as one chooses by a suitable choice of sub-intervals of the [0,1] interval (as determined by the choice of the integer n). Therefore, the lotteries in statements P.1 and Q.1 are “same-stakes” so long as y and zG are sufficiently close. This last inequality is satisfied by choosing z close to y/G, which is always possible because P.1 places no restriction on the size of z.

An example using numbers reported in Table 1 may help to explicate this point. Consider the three payoff amounts 14, 4, and 0 (used in one of the experiments run in Atlanta reported below). The theoretical point explained here about same stakes risk aversion is robust to all choices of payoff scale such as dollars, or dollars divided or multiplied by any power of 10. Clearly, 14/4 = 3.5 (the value of C for DTEU). Suppose that the value function υ for CPT is such that $\upsilon(14)/\upsilon(4) \geq 3$, then the value of C for CPT is at least 3. In that case, the entry in the first row and first column of Table 1 (for C = 3) tells us that 1,000 for sure is preferred to the 50/50 lottery that pays 33,000 or 0 for a DTEU or CPT agent. But the entries in Table 1 are derived from the negation of statement Q.1, with small adjustment of the positive payoff to get a strict preference $z \succ \{zG, 0.5\}$, where $G = 33$ if $n = 5$ and $C = 3$. To have the lotteries in statements P.1 and Q.1 be of the “same-stakes” set $z = 0.5$. Then $zG = 0.5 \times 33 = 16.5$ which is comparable to the high payoff of 14 in the P.1 example. An implication of this P.1 example is then seen to be that a sure payoff of 50 cents is preferred to the 50/50 lottery that pays $16.50 or 0. This is an example of implausible same-stakes risk aversion. It shows that DTEU and CPT are inconsistent with the following two plausible risk preferences holding simultaneously:

Q.1.e A 50/50 lottery that pays 16.50 or 0 is preferred to a sure payoff of 0.50; and

P.1.e A three-outcome lottery that pays 14, 4 or 0, with probabilities $p - 0.1$, 0.2 and $1 - p - 0.1$, is preferred to a two outcome lottery that pays 14 or 0, with probabilities p and $1 - p$, for all $p \in \{0.1, 0.2, ..., 0.8, 0.9\}$.

It is a straightforward exercise to verify that an expected utility of income model with CRRA preferences with r = 0.5 is consistent with both Q.1.e and P.1.e whereas CPT with estimated parameters such as those reported by Tversky and Kahneman (1992) is inconsistent with Q.1.e and P.1.e holding simultaneously.[4]

---

[4] At least 75% of the subjects in the Atlanta 14/4 experiment, reported in section VI, revealed preferences consistent with pattern P.1.e.

III. Calibrations with Varying Payoffs

Calibration propositions for theories with nonlinear utility of money payoffs have been reported in several papers (cited above in the introduction). In order to provide a foundation for our concavity calibration experiments, we report a calibration proposition for expected utility theory and a corollary that applies to rank-dependent theories. Design of experiments reported in section VII is based on the calibration patterns discussed here.

A. Calibration for Linear Probability Transformation Functions

We now revisit the large stakes risk aversion implications of postulated patterns of small stakes risk aversion for expected utility theory introduced into the literature by Rabin (2000). These implications hold for all three expected utility models discussed in Cox and Sadiraj (2006), the expected utility of terminal wealth model, the expected utility of income model, and the expected utility of initial wealth and income model. For bounded intervals of income, Proposition 2 states a concavity calibration result for expected utility theory with weakly concave utility of money payoff function $u(\cdot)$.[5] Let the variable x denote amounts of certain money payoff, interpreted either as initial wealth or exogenous income. Consider binary lotteries with gain amount g and loss amount ℓ that yield payoffs $x+g$ or $x-\ell$ to the agent.

Let $\lceil x \rceil$ denote the smallest integer larger than x and $f(\cdot)$ be the transformation function of decumulative probabilities. Define $A(r,K) = r^2 - r^{2+K} - r^K$ and $N_*(r) = 2 - \ln 2 / \ln(r)$.

Proposition 2. Let positive $\ell < g$ and integer $N > (1+g/\ell)N_*(\ell/g)$ be given. Denote

$M = N(g+\ell)$ and (*) $J = \lceil N - K + 1 + (\ell/g)^{-N} A(\ell/g, K) \rceil$, for integer $K \in (N_*(\ell/g), N)$.

Consider statements

P.2 $\quad x \succsim \{x+g, 0.5; x-\ell\}$, for all $x \in [m, M]$

Q.2 $\quad \{z + J(g+\ell), 0.5; m\} \succ z$, for $z = m + K(g+\ell)$, for some K.

a. Any DTEU agent who rejects $\{g, 0.5; -\ell\}$ satisfies P.2.
b. There are DTEU agents who satisfy both P.2 and Q.2.
c. There are no (u-concave) EU agents who satisfy both P.2 and Q.2.

Proof: See appendix A.1 and appendix A.3.

---

[5] See Rabin (2000) and Cox and Sadiraj (2006) for concavity calibrations on unbounded domains.

Note that: $K > 2 - \ln 2 / \ln(\ell / g)$ implies $A > 0$. Hence, for any given m and z, the fourth term on the right hand side of statement $(*)$ increases geometrically in M. This implies that for any amount of gain G, as big as one chooses, there exists a large enough interval in which preference for x over a risky lottery $\{x+g, 0.5; x-\ell\}$, for all integers x from the interval $[m, M]$, implies a preference for z for sure to the risky lottery $\{G, 0.5; m\}$. We use statements $(*)$ and Q.2 in Proposition 2 to construct the illustrative examples in Table 2.

Suppose that an agent prefers the certain amount of income x to the lottery $\{x+110, 0.5; x-100\}$, for all integers $x \in [100, M]$, where values of M are given in the "Rejection Intervals" column of Table 2. In that case all three expected utility (of terminal wealth, income, and initial wealth and income) models predict that the agent prefers receiving the amount of income 3,000 for sure to a risky lottery $\{G, 0.5; 100\}$, where the values of G are given in the first column of Table 2. For example, if $[m, M] = [100, 50000]$ then $G = 0.13 \times 10^{13}$ for all three expected utility models. According to the entry in the second column and M = 30,000 row of Table 2, expected utility theory predicts that if an agent prefers certain payoff in amount x to lottery $\{x+90, 0.5; x-50\}$, for all integers x between 100 and 30,000, then such an agent will prefer 3,000 for sure to the 50/50 lottery with positive outcomes of 100 or 0.13×10^{62}.

B. Calibration for Nonlinear Probability Transformation Functions

The following corollary to Proposition 2 applies to rank dependent theories including cumulative prospect theory (CPT) with zero-income reference point (Tversky and Kahneman, 1992). Define $r(t) = (1-t)\ell / tg$. Let $h(\cdot)$ denote the probability transformation function for the probability of the high payoff in a binary lottery. One has:

Corollary 2.1. Let $q = r(h(0.5))$ and integer $K > 2 - \ln 2 / \ln q$, be given. Let the value function be (weakly) concave on positive domain. There are no CPT agents who satisfy both P.2 and Q.2.

Proof: See appendix A.3.

Recall that for expected utility theory, with functional that is linear in probabilities, Proposition 2 reveals implausible large-stakes risk aversion if $g > \ell$. In the corollary, this implication holds when $h(0.5)g > [1-h(0.5)]\ell$. That is, for such lotteries, for any given m and z, the third term on the right hand side of inequality $(*)$ increases geometrically in M because

$q = r(h(0.5)) < 1$. This implies that if $h(0.5)g > [1 - h(0.5)]\ell$ then for any amount of gain G, as big as one chooses, there exists a large enough interval in which preference for x over a risky lottery $\{x + g, 0.5; x - \ell\}$, for all integers x from the interval $[m, M]$, implies a preference for z for sure to the risky lottery $\{G, 0.5; m\}$. Examples that illustrate the implications of Corollary 2.1 are similar to those in Table 2.

IV. Empirical Interpretation of Calibration Propositions

Previous calibration literature has been controversial. Some scholars (e.g., Rabin and Thaler, 2001; Wakker, 2005), appear to believe that it is obvious that individuals' risk preferences conform to the type of small-stakes risk aversion pattern supposed in Rabin's (2000) calibration. Others disagree, and argue that Rabin's supposed *small-stakes* risk aversion pattern is itself implausible, and that his calibration proposition has no empirical relevance. For example Watt (2002) noted, correctly, that Rabin's small-stakes risk aversion supposition is inconsistent with the Arrow-Pratt relative risk aversion measure being less than 170 for expected utility theory. He cites voluminous literature reporting empirical estimates of relative risk aversion measures with values much smaller than 170. Similar criticisms of the empirical relevance of Rabin's (2000) calibration proposition were stated by Palacios-Huerta and Serrano (2006).

The small-stakes risk aversion pattern used in our varying-probabilities calibration has elicited opinions of acceptance and rejection. Some economists find preference for the three-outcome lotteries supposed in our calibration to conform to their opinion. Others do not. And some have advanced critiques that follow the approach used by Watt and Palacios-Huerta and Serrano to critique Rabin's calibration. Some discussion of the empirical interpretation of calibration propositions seems warranted.

A. Interpretation of Proposition 2 and its Corollary

How does one interpret Proposition 2 and Corollary 2.1? First, they tell us that statements P.2 and Q.2 conform to the dual independence axiom that characterizes the dual theory of expected utility. Second, they tell us that statements P.2 and Q.2 are inconsistent for expected utility theory (EUT) and cumulative prospect theory (CPT) with zero-income reference point. Hence, these *theories predict*: if P.2 (the certain amounts x are preferred to the stated small-stakes lotteries) then Q.2 (large-stakes lotteries like those in Table 2 are rejected). Equivalently, the *theories predict* that if *not* Q.2 (large-stakes lotteries like those in Table 2 are *not* rejected) then *not* P.2 (the certain amounts *x* are *not* preferred to the small-stakes lotteries). There is general agreement that individuals would accept extremely favorable large-stakes lotteries like those in Table 2 *and* that theories of decision under risk are consistent with such acceptance. In that case, the calibration proposition and corollary tell us that EUT and CPT with zero-income reference point must be *inconsistent* with the small-stakes risk aversion pattern contained in statement P.2.[6] But this leaves us with an empirical question: will individuals actually reject the small-stakes lotteries in statement P.2, *as predicted by* EUT and CPT, or will they accept them? This is the empirical question addressed by the experiments we report in section VII.

B. Interpretation of Proposition 1 and its Corollaries

How does one interpret Proposition 1 and its corollaries? First, they tell us that statements P.1 and Q.1 conform to the independence axiom that characterizes expected utility theory. Second, they tell us that statements P.1 and Q.1 are inconsistent for the class of models NTDP that includes the dual theory of expected utility and variants of cumulative prospect theory with constant or variable reference amounts of payoff. Hence, these *theories predict*: if P.1 (the three-outcome lotteries are preferred to two-outcome lotteries then Q.1 (large-stakes lotteries like those in Table 1 are rejected). Equivalently, the *theories predict* that if *not* Q.1 (large-stakes lotteries like those in Table 1 are *not* rejected) then *not* P.1 (the three-outcome lotteries are *not* preferred to two-outcome lotteries). It appears that everyone agrees that individuals would accept extremely favorable large-stakes lotteries like those in Table 1 and that theories of decision under risk are consistent with such acceptance. In that case, the calibration proposition and corollaries tell us that dual theory and cumulative prospect theory must be *inconsistent* with the risk aversion pattern contained in statement P.1 in Proposition 1. Again, this leaves us with an empirical question: will individuals actually reject the two-outcome lotteries in P.1 in favor of the three-outcome lotteries, as predicted by dual theory and cumulative prospect theory, or will they choose the two-outcome lotteries? This is the empirical question addressed by the experiments we report in section VI.

---

[6] This statement for EUT is similar to Rabin and Thaler's (2002) response to Watt (2002).

V. Experiment Design Issues

We here discuss issues that arise in designing experiments with the calibration patterns contained in statements P.1 and P.2.

A. Power vs. Credibility with Varying-Probabilities Calibration Experiments

Table 1 illustrates the relationship between the ratio C of high payoff to intermediate payoff in the three-outcome lottery and the difference between probabilities in adjacent terms in the calibration (determined by the value of n in $\frac{i}{2n} - \frac{i-1}{2n}$). The design problem for varying-probability calibration experiments is inherent in the need to have small enough sub-intervals of the [0,1] interval for the calibration pattern in Proposition 1 to lead to the implication of clearly implausible risk aversion.

There are two problems with big values of the parameter n (i.e., large numbers of sub-intervals). First, a subject's decisions may involve trivial financial risk because the differences between all of the moments of the distributions of payoffs for the three-outcome lottery $\{y, p_i - \delta; x, 2\delta\}$ and the two-outcome lottery $\{y, p_i\}$ become insignificant as n becomes large. Consider, for example, the case of y = \$100, x = \$25, and $n = 500$. In this case, the difference between expected values of the two-outcome and three-outcome lotteries is 5 cents (for all *i*). The difference between standard deviations of payoffs for the two-outcome and three-outcome lotteries, at *i* = 500, is 4 cents.

The second problem with large *n* is that adjacent probabilities differ by only $1/2n$ while the subject's decision task is to make $2n-1$ choices. For example, for $n = 500$ adjacent probabilities differ by 0.001 and the subjects' decision task is to make 999 choices. In such a case, the subjects may not be sensitive to the probability differences and the payoffs may not dominate decision costs because of the huge number of choices needing to be made. In contrast, if the length of each sub-interval is $1/10$ (i.e. $n = 5$) then the difference in expected payoffs between the two-outcome and three-outcome lotteries is \$5 for the above values $y = \$100$ and $x = \$25$, and for *i* = 5 the difference in standard deviations is \$4.17; furthermore, the subjects' decision task is to make only 9 choices. The calibration implications of *n* = 5 are less spectacular than for *n* = 500, as shown in Table 1, but the resulting experiment can credibly be implemented. In our experiments, we use relatively low values of the parameter *n*.

B. Affordability vs. Credibility with Varying-Payoffs Calibration Experiments

Table 2 illustrates the relationship between the size of the interval $[m, M]$ in the left-most column, used in the supposition underlying a utility of money payoff calibration, and the size of the high gain G in the result reported in the other columns of the table. Varying-payoff calibration experiments involve tradeoffs between what is affordable and what is credible, as we shall next explain.

As an example, consider an experiment in which subjects were asked to choose between $\$x$ for sure and the binary lottery $\{\$x+\$110, 0.5; \$x-\$100\}$ for all x between m = \$1,000 and M = \$350,000. Suppose the subject always chooses the certain amount $\$x$ and that one of the subject's decisions is randomly selected for payoff. Then the expected payoff to a single subject would exceed \$175,000. With a sample size of 30 subjects, the expected payoff to subjects would exceed \$5 million, which would clearly be unaffordable. But why use payoffs denominated in U.S. dollars? After all, Proposition 2 is dimension invariant. Thus, instead of interpreting the figures in Table 2 as dollars, they could be interpreted as dollars divided by 10,000; in that case the example experiment would cost about \$500 for subject payments and clearly be affordable. So what is the source of the difficulty? The source of the difficult tradeoff for experiment design becomes clear from inspection of Proposition 2: the unit of measure for m and M is the same as that for the loss and gain amounts ℓ and g in the binary lotteries. If the unit of measure for m and M is \$1/10,000 then the unit of measure for ℓ and g is the same (or else the calibration doesn't apply); in that case the binary lottery has high and low payoffs in amounts \$0.0001x + \$0.011 and \$0.0001x - \$0.010, which involves trivial financial risk of 2.1 cents.

The design problem for concavity calibration experiments with money payoffs is inherent in the need to calibrate over an $[m, M]$ interval of sufficient length for the calibrations in Proposition 2 and Corollary 2.1 to lead to the implication of implausible risk aversion in the large. There is no way to avoid this problem; the design of any varying-payoffs calibration experiment will reflect a tradeoff between affordability of the payoffs and credibility of the incentives. In our experiments, we address this problem in two ways by: (a) conducting some experiments in India, where we can afford to use $[m, M]$ intervals of rupee payoffs that are sufficiently wide for calibration to have bite; and (b) conducting an experiment in Germany, partly on the floor of a casino, which makes use of large contingent euro payoffs affordable.

VI. Experiments with Varying Probabilities

We ran four varying-probabilities calibration experiments in Germany, India, and the United States. We explain the common design features and idiosyncratic lotteries in these experiments and present a more detailed discussion of one experiment to provide a representative example. We begin with the example.

A. Experiment Design: An Example

Subjects in one experiment parameterization were asked to make choices for each of the nine pairs of lotteries shown in Table 3. The fractions in the rows of the table are the probabilities of receiving the prizes in the two outcome (option A) and three outcome (option B) lotteries. Each row of Table 3 shows a pair of lotteries included in the experiment. The subjects were *not* presented with a fixed order of lottery pairs, as in Table 3. Instead, each lottery pair was shown on a separate (response form) page. Each subject picked up a set of response pages that were arranged in independently drawn random order. He or she could mark choices in any order desired. On each decision page, a subject was asked to choose among a two outcome lottery (option A in some row of Table 3), a three outcome lottery (option B in the same row of Table 3), and indifference ("option I").

B. Experiment Design: Alternative Parameterizations and Protocols

We conducted four experiments on empirical validity of the calibration pattern P.1 postulated in Proposition 1. One experiment parameterization uses pairs of two outcome and three outcome lotteries $A_j = \{y, p_j\}$, and $B_j = \{y, p_j - 0.1; x, 0.2\}$, for $j = 1, 2, \cdots 9$, and $y = 14$, $x = 4$ as shown in Table 3. We also ran experiments with the parameterizations $(y, x) = (40, 10)$ and (400, 80).

The experiments were conducted in Magdeburg (Germany), Atlanta (U.S.A.) and Calcutta (India) with payoffs, respectively, in euros, U.S. dollars, and Indian rupees. The experiments used the following parameters: Magdeburg 40/10: $y =$ 40 euros, $x =$ 10 euros. Atlanta 40/10: $y =$ 40 dollars, $x =$ 10 dollars. Atlanta 14/4: $y = 14$ dollars, $x =$ 4 dollars. Calcutta 400/80: $y =$ 400 rupees, $x =$ 80 rupees. Economic significance of the rupee payoffs is discussed in section VII.C. The payoff protocol used random selection of one decision for payoff, which is a standard procedure used in testing theories of decision under risk with or without the independence axiom. Experimental tests of random selection have generally reported consistency with the isolation effect of subjects focusing on individual decision tasks (Camerer, 1989; Starmer and Sugden, 1991; Beattie and Loomes, 1997; Cubitt, Starmer, and Sugden, 1998; Hey and Lee, 2005a, b; Laury, 2006; Lee, 2008). An appendix available from the authors reports subject instructions (in English), response forms (or pages), and detailed information on the protocol used in all of the experiments.

C. Data Provide Support for Calibration Pattern P.1

In testing for the presence of choices that satisfy the calibration pattern, we aggregate choices of option B with (the very small number of) choices of option I (indifference) because statement P.1 in Proposition 1 involves weak preference for B over A. Aggregated choices of B and I are reported as B^I. Subjects' choice patterns are recorded as sequences of nine letters, ordered according to the probability of the high outcome. For example, the pattern [A, B^I, B^I, A, B^I, B^I, B^I, B^I, A] would indicate that a subject chose A (a two outcome lottery) when the probability of the high outcome was 1/10, 4/10 and 9/10 - indexed as $j = 1$, 4, and 9 - and chose B or I for all other values of the index j. For the experiment with the parameterization as shown in Table 3, this choice pattern would mean the subject chose option A on (randomly ordered) pages with the lottery pairs in rows 1, 4, and 9 in the table and chose option B or option I on all other pages.

We use error-rate analysis for statistical inferences on the proportion of subjects who made choices consistent with the calibration patterns.[7] Choice probabilities are assumed to deviate from 1 or 0 by an error rate ε, as in Harless and Camerer (1994). Thus if B^I is preferred to A then Prob(choose B^I) = $1-\varepsilon$ and if B^I is *not* preferred to A then Prob(choose B^I) = ε, where $\varepsilon < 0.5$. The error rate model postulates that a subject with real preferences for B^I (respectively A) over A (respectively B^I) in all nine lottery pairs could nevertheless be observed to have chosen the other option in some rows. For example, according to this model a subject with underlying preferences [B^I, B^I, B^I, B^I, B^I, B^I, B^I, B^I, B^I] could, instead, be observed to

[7] We are grateful to Nathaniel Wilcox for generous advice about this approach to data analysis and for supplying SAS code. See Wilcox (2008) for discussion of econometric methods for analysis of data from binary discrete choice under risk.

choose a different pattern such as [B^I, B^I, A, B^I, A, B^I, B^I, B^I, B^I], an event with probability $(1-\varepsilon)^7\varepsilon^2$.

Stochastic choice Model I contains only the choice pattern with a sequence of nine B^I in the category "calibration pattern" and its dual ("mirror") image with a sequence of nine A in the "other pattern." According to Proposition 1, this calibration pattern implies that 1,000 for sure is preferred to the 50/50 lottery that pays 98,000 or 0 for the Atlanta 14/4 experiment, as reported in the top-most of the shaded rows in Table 4. For the Calcutta 400/80 experiment, Proposition 1 implies that 1,000 for sure is preferred to the 50/50 lottery that pays 1 million or 0, as reported in the shaded row for the Calcutta 400/80 listing in Table 4.

Model I is overly conservative in its specification of calibration patterns because other data patterns can be calibrated to imply implausible risk aversion. Stochastic choice Model II includes two patterns in the category "calibration patterns": the pattern with choice of B^I for index $j = 1, 2, \cdots, 8$ and the all B^I pattern (that is, $j = 1, 2, \cdots, 9$). The mirror images of these two patterns comprise the "other patterns" for Model II. Application of Corollary 1.3 demonstrates that these two calibration patterns of "no A except for index $j = 9$" imply that 1,000 for sure is preferred to the 50/50 lottery that pays 81,000 or 0, as reported for the Atlanta 40/10 experiment listings in Table 4. We also consider Model III which includes the patterns "no A except for indexes $j = 8$ and/or 9" in the category of calibration patterns. The mirror images of these four patterns comprise the other patterns for Model III. An implication of Corollary 1.3 for these four calibration patterns in case of $n = 5$ and C = 4 is preference for 1,000 for sure to the 50/50 lottery that pays 27,000 or 0, as shown in the Atlanta 40/10 and Magdeburg 40/10 listings in the table.

Table 4 reports results from maximum likelihood estimation of the proportion of subjects who exhibit the calibration patterns for Models I, II and III. Estimations are reported for a single error rate for all choices, for two different error rates (one error rate for choices with index j =1,..,4 and another one for choices with index j =5,..,9), and three different error rates (one error rate for choices with index $j = 1, 2, 3$, another error rate for choices with index $j = 4, 5, 6$, and another one for choices with index $j = 7, 8, 9$).[8]

---

[8] The three error rate models can capture subjects' different sensitivities to high, intermediate and low probabilities of the high outcome.

The first row of Table 4 shows results for the Atlanta 14/4 experiment data. For Model I with one error rate the estimated proportion of subjects who exhibited the calibration pattern is 0.74. The Wald 90 percent confidence interval is (0.55, 0.93). The 0.74 estimate is significant at one percent (as indicated by **). The other columns in the first row of Table 4 report the estimated proportions of subjects whose choice patterns in Atlanta 14/4 conform to calibration patterns with the 1 error, 2 error, and 3 error rate versions of Models I, II, and III. These estimates vary between 0.74 and 0.90, and they are all significant at one percent. The entries in bold font indicate the model that is selected by likelihood ratio tests; that is, with data from Atlanta 14/4, Model I with 1 error or 3 errors and Models II and III with 1 error, 2 errors, or 3 errors are all rejected in favor of Model I with 2 error rates.

The second through fourth rows of Table 4 show the estimated proportions of subjects whose choices are consistent with calibration patterns in experiments Atlanta 40/10, Magdeburg 40/10, and Calcutta 400/80. Depending on the model and number of errors, the estimated proportion of subjects with data consistent with the calibration patterns in Atlanta 40/10 varies from 0.56 to 0.63, all significant at one percent. The estimates for data from Magdeburg 40/10 vary from 0.37 to 0.41, all significant at one percent. Estimates with data from experiment Calcutta 400/80 lie between 0.72 and 0.74, all are significant at one percent. The entries in bold font indicate the model that is selected by likelihood ratio tests over all other models in that row.

VII. Experiments with Varying Payoffs

We ran three experiments with calibration patterns for payoff transformation theories identified in Proposition 2 and Corollary 2 in India and Germany. We explain the common features and idiosyncratic lotteries used in these experiments after presenting a detailed discussion of one experiment to provide a representative example.

A. Experiment Design: An Example

Subjects in one experiment parameterization were asked to make six choices between a certain amount of money x and a binary lottery $\{x+30, 0.5; x-20\}$ for values of x from the set {100, 1K, 2K, 4K, 5K, 6K}, where K = 1,000. Subjects were asked to choose among option A (the risky lottery), option B (the certain amount of money), and option I (indifference). The choice tasks given to the subjects for this parameterization are presented in Table 5. Each row of Table 5 shows a certain amount of money and paired lottery in a choice task included in the experiment. The subjects were *not* presented with a fixed order of decision tasks, as in Table 5.

Instead, each pair of sure payoff and lottery was shown on a separate (response form) page. Each subject picked up a set of response pages that were arranged in independently drawn random order. He or she could mark choices in any order desired.

B. Experiment Design: Alternative Parameterizations and Protocols

We conducted three experiments on empirical validity of the calibration pattern P.2 in Proposition 2. These experiments used the random decision selection payoff protocol. Calcutta $30/-20$: binary lotteries $\{x+30, 0.5; x-20\}$ and sure payoffs x from the set {100, 1K, 2K, 4K, 5K, 6K}, where K = 1,000; payoffs in rupees. Calcutta $90/-50$: binary lotteries $\{x+90, 0.5; x-50\}$ for values of x from the set {50, 800, 1.7K, 2.7K, 3.8K, 5K}, where K = 1,000; payoffs in rupees. Magdeburg $110/-100$: binary lotteries $\{x+110, 0.5; x-100\}$ for values of x from the set {3*K*, 9*K*, 50*K*, 70*K*, 90*K*, 110*K*}, where K = 1,000; payoffs in *contingent* euros.

An appendix available from the authors reports the subject instructions (in English), the response forms (or pages), and detailed information on the protocol used in all of the experiments. Before presenting data, we discuss economic significance of the rupee payoffs in Calcutta experiments and the meaning of *contingent* euro payoffs in the Magdeburg experiment.

C. Economic Significance of the Rupee Payoffs

At the time of the first experiment in Calcutta (2004), data collected show most student subjects' incomes included only scholarships that paid stipends of 1,200-1,500 rupees per month in addition to the standard tuition waiver that each received. This means that the highest certain payoff used in the Calcutta $30/-20$ experiment (6,000 rupees) was equal to four or five months' stipend for the subjects. The daily rate of pay for the students was 40 to 50 rupees. Hence the amount at risk in the Calcutta $30/-20$ experiment lotteries (the difference between the high and low payoffs) was greater than or equal to a full day's pay. The amount at risk in the Calcutta $90/-50$ experiment (140 rupees) was almost three times as large.

A sample of commodity prices in Calcutta at the time of the first experiment conducted there (Calcutta $30/-20$) is reported in an appendix available from the authors. Prices of food items were reported in number of rupees per kilogram. There are about 15 servings in a kilogram of these food items.[9] As reported in the appendix table, for example, we observed prices for poultry of 45 – 50 rupees per kilogram. Hence, the size of the risk in the lotteries in Calcutta $30/-20$ (50 rupees) was equivalent to 15 servings of poultry. The price of a moderate quality restaurant meal was 15 – 35 rupees per person. Hence the 50 rupee risk in the experiment lotteries was the equivalent of about 1.5 – 3 moderate quality restaurant meals. The observed prices for local bus tickets were 3 – 4.5 rupees per ticket. This implies that the 50 rupee risk in the experiment lotteries was the equivalent of about 14 bus tickets. Again, the amount at risk in the Calcutta $90/-50$ experiment was about three times as large.

D. Contingent Euro Payoffs in Magdeburg

The Magdeburg $110/-100$ experiment included amounts x that were as large as 110K euros. We could credibly offer to pay such large amounts in *contingent* euros by using the following protocol. The experiment included two parts. In part 1 subjects made their choices between the sure amounts and the lotteries in the MAX-Lab at the University of Magdeburg. They were told that whether their payoffs would be hypothetical or real depended on a condition which would be described later in part 2. After making their decisions the subjects were informed that real payoffs were conditional on winning gambles at the Magdeburg Casino. The payoff contingency was implemented in the following way. For each participant the experimenter placed €90 on the number 19 on one of the (four American) roulette wheels at the casino. The probability that this bet wins is 1/38. If the bet wins, it pays 35 to 1. If the first bet won, then the experimenter would bet all of the winnings on the number 23. If both the first and second bet won, then the payoff would be €(35 $\times$ 35 $\times$ 90) = €110,250, which would provide enough money to make it feasible to pay any of the amounts involved in the part 1 decision tasks for that subject. The real payoff contingency was made credible to the subjects by randomly selecting three of them to accompany the experimenter to the casino and subsequently report to the others whether the experimenter had correctly placed the bets and recorded the outcomes.

E. Data Provide Some Support for the Concavity Calibration Pattern P.2

Statement P.2 in Proposition 2 involves weak preference for option B over option A. Therefore, in all tests we aggregate choices of option B with (the very small number of) choices

---

[9] There are 2.205 pounds per kilogram and 16 ounces in a pound, hence there are 35.28 ounces per kilogram. The U.S. Department of Agriculture's food pyramid guide defines a "serving" of meat, poultry, or fish as consisting of 2 – 3 ounces.

of option I (indifference) and denote the aggregated choice category as B^I. We report tests for incidence in the data of patterns of choices that, according to Proposition 2 and Corollary 2.1, imply implausible risk aversion in the large with expected utility theory and, for two of the experiments, with original cumulative prospect theory (with zero-income reference point) and with rank dependent expected utility theory.

We use error rate models to draw statistical conclusions from these data. Recall that this type of analysis takes into account that a subject with real preferences for option B^I rather than option A in all six rows could nevertheless be observed to have chosen B^I in five (or fewer) out of six rows. That is, the model assumes that a subject with real underlying preferences such as [B^I, B^I, B^I, B^I, B^I, B^I] could, instead, choose a different pattern, say [B^I, B^I, B^I, A, B^I, B^I], an event with probability $(1-\varepsilon)^5\varepsilon$, where ε is an error rate.

Models I, II, and III considered here are as follows. Model I includes only choices of all B^I (corresponding to M = 6,000 in Proposition 2 for the Calcutta 30/-20 experiment for example) as a calibration pattern and its mirror, all A's as the other pattern. Let the small stakes lotteries be $\{x+30, 0.5; x\text{-}20\}$ for x from 100 to 6,000. According to Proposition 2, the choice pattern "all B^I" implies that 1,000 for sure is preferred to the lottery that pays $0.39\text{x}10^{23}$ or 0 with equal probabilities. Model II (which corresponds to Proposition 2 with M = 5,000 for the Calcutta 30/-20 experiment) contains the Model I pair of (calibration and other) patterns, and one additional calibration pattern with A as the last entry (for x = 6,000) and its mirror image as an additional "other pattern." According to Proposition 2, the calibration patterns in Model II imply that getting 1,000 for sure is preferred to the 50/50 lottery that pays 0.12×10^{20} or 0. Finally, Model III (which corresponds to Proposition 2 with M = 4,000 for the Calcutta 30/-20 experiment) contains patterns with four sequential B^I in the first four positions (for x = 100, 1000, 2000, and 4000) as calibration patterns and their mirror images as other patterns. With these calibration patterns, Proposition 2 implies that getting 1,000 for sure is preferred to the lottery that pays 0.36×10^{16} or 0 with equal probabilities.

The top row in Table 6 shows estimated proportions of subjects whose choices satisfy the calibration patterns with the 1 error, 2 error, and 3 error rate versions of Models I, II, and III using data from Calcutta 90/-50. The estimated proportions for the 1 error rate version of Model I (M = 5,000) is 0.82, with Wald 90 percent confidence interval (0.70, 0.94). The estimated proportions for all models vary between 0.80 and 0.82; all are significant at one percent (indicated by **). The bold entries indicate the models that are selected by likelihood ratio tests.

The second row of Table 6 reports estimates for data from Calcutta 30/-20. The estimated proportions vary between 0.36 and 0.48, and all are significant at one percent. The estimations for Calcutta 30/-20 imply that 36 to 48 percent of the subjects in this experiment made choices that conform to calibration patterns for expected utility theory, rank dependent utility theory, and original cumulative prospect theory. Estimates in the third row for data from Magdeburg 110/-100 vary between 0.50 and 0.56; all are significant at one percent.

VIII. Is There a Plausible Decision Theory for Risky Environments?

The expected utility of terminal wealth model provides a complete theory of rational decisions under risk. Classic papers by Arrow (1971) and Pratt (1964) develop this theory for risks of all scales by defining the domain of discourse as the real line. They assume bounded utility, presumably to avoid exposing the theory to the generalized St. Petersburg paradox. But bounded utility exposes expected utility theory and other popular theories of decision under risk to implausible risk aversion. As shown by Cox and Sadiraj (2008), with an unbounded domain expected utility theory, cumulative prospect theory, rank dependent expected utility theory, and the dual theory of expected utility exhibit either a generalized St. Petersburg paradox (with unbounded utility) or implausible risk aversion (with bounded utility). Bounding the domain of the theory would seem to provide a solution to this problem but the potential for implausible risk aversion remains, as shown by the dual calibration propositions and corollaries reported herein.

Prominent theories of decision under risk model individuals' preferences over lotteries with nonlinear transformation of money payoffs and/or nonlinear transformation of probabilities. Previous calibration literature has focused on the implications of nonlinear transformation of money payoffs. This paper introduces a dual calibration that focuses on the implications of nonlinear transformation of probabilities. Theories with functionals that are nonlinear in both probabilities and payoffs are vulnerable to both calibration patterns unless they incorporate variable reference amounts of payoff. With variable reference amounts of payoff, the dual calibration for nonlinear probability transformations introduced herein is problematic for these theories but the generalized Rabin pattern applied elsewhere in the literature is not.

The two calibration propositions provide a paradoxical insight into theories of risk aversion in that certain patterns of risk aversion that conform to the independence axiom (respectively, dual independence axiom) imply implausible large stakes risk aversion for the dual theory of expected utility (respectively, expected utility theory). As such, a pattern of risk aversion that characterizes rational behavior for a theory with utility functional that is linear in probabilities (respectively, linear in payoffs) has implausible implications for a theory with functional that is linear in payoffs (respectively, linear in probabilities). Corollaries to the propositions show that theories with functionals that are nonlinear in both payoffs and probabilities have problems that can be demonstrated with either the varying-payoffs pattern of risk aversion reported in previous literature or the varying-probabilities pattern reported herein. The varying-payoffs pattern of small-stakes risk aversion reported in previous literature is inconsistent with plausible large-stakes risk aversion for theories that incorporate decreasing marginal utility of money. The varying-probabilities pattern of risk aversion introduced herein reveals that theories with nonlinear transformation of probabilities have problems with implausibility of *same*-stakes risk aversion as well as small-stakes vs. large-stakes problems.

Previous literature has offered no real-payoff, controlled experiment data on the empirical relevance of patterns of risk aversion that have calibration implications. This paper reports data from seven experiments. The data provide some support for the empirical validity of risk aversion patterns underlying both of the dual calibrations.

The calibration propositions and corollaries suggest a central question: What type of model of risk-avoiding preferences is *not* called into question by the dual calibration critique? Here are the properties of a simple model that is immune to calibration. The utility functional for the model: (a) is linear in probabilities; and (b) has a variable reference point. This simple model is immune to calibration with the varying-probabilities pattern P.1 because of its linearity in probabilities. It is immune to calibration with the varying-payoffs pattern P.2 because its variable reference point can be identified with the certain amount x in each paired choice. Although this simple model survives the dual calibration critique, it has testable implications that can be easily called into question by data from experiments. Some more complicated variable reference point models that are linear in probabilities may have more empirical validity than this simple model.

References

Arrow, Kenneth J., *Essays in the Theory of Risk-Bearing*. Markham, Chicago, 1971.

Barberis, Nicholas, Ming Huang, and Richard H. Thaler, "Individual Preferences, Monetary Gambles, and Stock Market Participation: A Case for Narrow Framing," *American Economic Review*, 2006, 96(4), 1069-1090.

Beattie, Jane and Graham Loomes, "The Impact of Incentives Upon Risky Choice Experiments," *Journal of Risk and Uncertainty*, 1997, 2, 155-168.

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

Cox, James C. and Vjollca Sadiraj, "Small- and Large-Stakes Risk Aversion: Implications of Concavity Calibration for Decision Theory," *Games and Economic Behavior*, 2006, 56 (1), 45-60.

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

Cubitt, Robin, Chris Starmer and Robert Sudgen, "On the Validity of the Random Lottery Incentive System," *Experimental Economics*, 1998, 1, 115-31.

Harless, David, and Colin F. Camerer, "The Predictive Utility of Generalized Expected Utility Theories," *Econometrica*, 1994, 62, 1251-90.

Hey, John D. and Jinkwon Lee, "Do Subjects Separate (or Are They Sophisticated)?," *Experimental Economics*, 2005a, 8, 233-265.

Hey, John D. and Jinkwon Lee, "Do Subjects Remember the Past?," *Applied Economics*, 2005b, 37, 9-18.

Laury, Susan K., "Pay One or Pay All: Random Selection of One Choice for Payment," E*x*CEN Working Paper 2006-24, Georgia State University 2006.

Lee, Jinkwon, "The effect of the Background Risk in a Simple Chance Improving Decision Model," *Journal of Risk and Uncertainty, 2008, 36, 19-41.*

Neilson, William S., "Calibration Results for Rank-Dependent Expected Utility," *Economics Bulletin,* 2001, 4, 1-5.

Palacios-Huerta, Ignacio and Roberto Serrano, "Rejecting Small Gambles under Expected Utility," *Economic Letters*, 2006, 91, 250-259.

Pratt, John W., "Risk Aversion in the Small and in the Large," *Econometrica*, 32, 1964, 122-136.

Quiggin, John, *Generalized Expected Utility Theory. The Rank-Dependent Model*. Boston: Kluwer Academic Publishers, 1993.

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

Rabin, Matthew and Richard H. Thaler, "Anomalies: Risk Aversion," *Journal of Economic Perspectives*, 2001, 15, 219-232.

Rabin, Matthew and Richard H. Thaler, "Response from Matthew Rabin and Richard H. Thaler," *Journal of Economic Perspectives*, 2002, 16, 229-230.

Rubinstein, Ariel, "Dilemmas of an Economic Theorist," *Econometrica*, 2006, 74, 865-883.

Safra, Zvi and Uzi Segal, "Calibration Results for Non-Expected Utility Theories," *Econometrica*, 2008, 76, 1143-1166.

Starmer, Chris and Robert Sugden, "Does the Random-Lottery Incentive System Elicit True Preferences?," *American Economic Review*, 1991, 81, 971-79.

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

Wakker, Peter P., "Formalizing Reference Dependence and Initial Wealth in Rabin's Calibration Theorem," Econometric Institute, Erasmus University, Rotterdam, 2005.

Watt, Richard. "Defending Expected Utility Theory," *Journal of Economic Perspectives*, 2002, 16, 227-228.

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

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

Table 1. Calibrations for Varying-Probabilities Patterns

$$1000 \succ \{\overline{G}, 0.5\}$$

Rejection Sub-intervals	Calibration for C = 2.5	Calibration for C = 3	Calibration for C = 3.5*	Calibration for C = 4.0*	Calibration for C = 5.0*
N	$\overline{G}$	$\overline{G}$	$\overline{G}$	$\overline{G}$	$\overline{G}$
5	8,593	33,000	98,000	244,000	1,025,000
10	58,665	1,025,000	9,530,000	0.59×10^{8}	0.10×10^{10}
20	3,326,256	0.10×10^{10}	0.90×10^{11}	0.34×10^{13}	0.10×10^{16}
50	0.63×10^{12}	0.11×10^{19}	0.78×10^{23}	0.71×10^{27}	0.12×10^{34}
100	0.40×10^{21}	0.12×10^{34}	0.62×10^{43}	0.51×10^{51}	0.16×10^{64}
200	0.16×10^{39}	0.16×10^{64}	0.38×10^{83}	0.26×10^{99}	0.25×10^{124}
500	0.11×10^{92}	0.32×10^{154}	0.93×10^{202}	0.36×10^{242}	0.10×10^{305}

* $C = \frac{y}{x}, \frac{\upsilon(y)}{\upsilon(x)}, \frac{\upsilon(y-x)}{-\mu(-x)}$; *y/x* = 3.5 for Atlanta 14/4; *y/x*=4 for Magdeburg 40/10 and Atlanta 40/10; *y/x*=5 for Calcutta 400/80.

Table 2. Calibrations for Varying-Payoffs Patterns

$$3,000 \succ \{G, 0.5; 100\}$$

Rejection Intervals [100, *M*]	**Calibration for g=110, ℓ=100**	**Calibration for g=90, ℓ=50**	**Calibration for g=30, ℓ=20**
M	*G*	*G*	*G*
5,000	7,000	0.47×10^{16}	0.65×10^{29}
6,000	9,000	0.29×10^{18}	0.21×10^{33}
8,000	15,000	0.10×10^{22}	0.24×10^{40}
10,000	29,000	0.40×10^{25}	0.26×10^{47}
30,000	0.16×10^{9}	0.13×10^{62}	0.72×10^{117}
50,000	0.13×10^{13}	0.41×10^{98}	0.19×10^{188}

Table 3. Choice Alternatives in Varying-Probabilities Experiment Atlanta 14/4

Row	**Option A**		**Option B**		
	Payoff 14	Payoff 0	Payoff 14	Payoff 4	Payoff 0
1	1/10	9/10	0/10	2/10	8/10
2	2/10	8/10	1/10	2/10	7/10
3	3/10	7/10	2/10	2/10	6/10
4	4/10	6/10	3/10	2/10	5/10
5	5/10	5/10	4/10	2/10	4/10
6	6/10	4/10	5/10	2/10	3/10
7	7/10	3/10	6/10	2/10	2/10
8	8/10	2/10	7/10	2/10	1/10
9	9/10	1/10	8/10	2/10	0/10

Table 4. Error Rate Models and Predictions for Varying-Probabilities Experiments

Experiment	Nr. Of Subjects	Model I			Model II			Model III		
		1 error	2 errors	3errors	1 error	2 errors	3 errors	1 error	2 errors	3 errors
Atlanta 14/4	39	0.74** (.55,.93)	**0.90** (.81,.98)**	0.90** (.81,.99)	0.82** (.68,.96)	0.90** (.71,1.0)	0.90** (.77,1.0)	0.88** (.77,.99)	0.90** (.80,.99)	0.90** (.80,.99)
		1000 $\succ$ {98000, 0.5;0}			1000 $\succ$ {39000, 0.5;0}			1000 $\succ$ {15700, 0.5;0}		
Atlanta 40/10	22	0.56** (.37,.75)	0.62** (.42,.82)	0.62** (.34,.90)	0.59** (.39,.78)	**0.63** (.42,.83)**	0.63** (.41,.85)	0.60** (.39,.80)	0.61** (.38,.83)	0.60** (.35,.85)
		1000 $\succ$ {244000, 0.5;0}			1000 $\succ$ {81000, 0.5;0}			1000 $\succ$ {27000, 0.5;0}		
Magdeburg 40/10	31	0.38** (.20,.56)	**0.40** (.22,.59)**	0.41** (.22,.61)	0.37** (.19,.55)	0.39** (.21,.57)	0.40** (.21,.60)	0.40** (.21,.58)	0.40** (.22,.57)	0.41** (.22, .60)
		1000 $\succ$ {244000, 0.5;0}			1000 $\succ$ {81000, 0.5;0}			1000 $\succ$ {27000, 0.5;0}		
Calcutta 400/80	40	0.72** (.58,.86)	0.74** (.60,.88)	**0.74** (.60,.87)**	0.72** (.58,.86)	0.73** (.58,.88)	0.74** (.59,.88)	0.73** (.59,.86)	0.72** (.57,.87)	0.73** (.58,.87)
		1000 $\succ$ {1 million, 0.5;0}			1000 $\succ$ {256000, 0.5;0}			1000 $\succ$ {64000, 0.5;0}		

Table 5. Choice Alternatives in Varying-Payoffs Experiment Calcutta 30/-20

Row	Option A	Option B
1	80 or 130	100
2	980 or 1,030	1,000
3	1,980 or 2,030	2,000
4	3,980 or 4,030	4,000
5	4,980 or 5,030	5,000
6	5,980 or 6,030	6,000

Table 6. Error Rate Models and Predictions for Varying-Payoffs Experiments

Experiment	Nr. of subjects	Model I			Model II			Model III		
		1 error	2 errors	3errors	1 error	2 errors	3 errors	1 error	2 errors	3 errors
Calcutta 90/-50 m = 50	40	**0.82** (.70,.94)**	0.81** (.69,.93)	0.81** .68,.94	0.81** (.69,.93)	0.80** (.67,.93)	0.80** (.66,.94)	0.82** (.69,.94)	0.80** (.68,.93)	0.81** (.67,.95)
		M=5,000: 1000 ≻ {0.11x10^{13}, 0.5; 50}			M=4,000: 1,000 ≻ {0.18x10^{11}, 0.5; 50}			M=3,000: 1,000 ≻ {0.29x10^{9}, 0.5; 50}		
Calcutta 30/-20 m = 100	30	0.36* (.14,.59)	0.43** (.25,.62)	0.44** (.25,.64)	0.48** (.20,0.53)	0.43** (.17,.68)	0.46** (.27,.65)	0.48** (.30,.67)	0.37** (.20,.53)	**0.47** (.26,.68)**
		M=6,000: 1,000 ≻ {0.19x10^{26}, 0.5; 100}			M=5,000: 1,000 ≻ {0.59x10^{22}, 0.5; 100}			M=4,000: 1,000 ≻ {0.17 x10^{19}, 0.5;100}		
Magdeburg 110/-100 m = 3000	42	0.54** (.39,.68)	0.55** (.41,.68)	**0.54** (.40,.68)**	0.54** (.39,.68)	0.56** (.43,.69)	0.50** (.41,.68)	0.50** (.32,.68)	0.52** (.38,.66)	0.50** (.36,.64)
		M = 110,000: 5,000 ≻ {0.30x10^{23}, 0.5;3,000}			M= 90,000: 5,000 ≻ {0.35x10^{19}, 0.5; 3,000}			M= 70,000: 5,000 ≻ {0.41x10^{15}, 0.5;3,000}		

Appendix

A.1. Concavity Calibration Pattern and the Dual Independence Axiom

Let Γ denote the set of all decumulative distribution functions. Let $\oplus$ denote the following operator: $\lambda G \oplus (1-\lambda)H = (\lambda G^{-1} + (1-\lambda)H^{-1})^{-1}$, $\forall G, H \in \Gamma$. The dual independence axiom as stated in Yaari 1987 (p. 99) is:

Axiom DI: If G, G', and H belong to Γ and α is a real number satisfying $0 \le \alpha \le 1$, then $G \succeq G'$ implies $\alpha G \oplus (1-\alpha)H \succeq \alpha G' \oplus (1-\alpha)H$.

Suppose that a dual expected utility agent rejects binary lottery $\{x+g, 0.5; x-\ell\}$ in favor of receiving x for sure for some $x > 0$. Then by axiom DI and continuity the agent rejects $\{y+g, 0.5; y-\ell\}$ in favor of getting y for sure for all positive y .

Let F_x denote the decumulative distribution function for the binary lottery and D_x the decumulative distribution for the degenerate lottery that pays x for sure. Then the agent's preference for the sure amount x over the binary lottery $\{x+g, 0.5; x-\ell\}$ is formally written as

(∗) $D_x \succeq F_x$.

Let y be given. Without any loss of generality assume that y > x. Then take a sequence of $\alpha_n \in (0,1)$, $n \in N$ such that $\alpha_n \underset{n\to\infty}{\to} 1$. For each $\alpha_n, n \in N$ there exists a $z_n \ge y$ such that $y = \alpha_n x + (1-\alpha_n) z_n$. Let D_{z_α} denote the decumulative distribution for the degenerate lottery that pays z_n for sure. Statement (∗) and Axiom DI imply $\alpha D_x \oplus (1-\alpha) D_{z_\alpha} \succeq \alpha F_x \oplus (1-\alpha) D_{z_\alpha}$. Note that by definition of operator $\oplus$ and the construction of z_n, the expression on the left hand side is the degenerate lottery that pays y for sure whereas the one on the right is the binary lottery $\{y+\alpha_n g, 0.5; y-\alpha_n \ell\}$. So, the agent prefers getting y for sure to a 50/50 lottery with payoffs $y+\alpha_n g$ or $y-\alpha_n \ell$, for all α_n . By continuity our agent (weakly) prefers getting y for sure to the binary lottery $\{y+g, 0.5; y-\ell\}$.

A.2. Proof of Proposition 1 and its Corollaries

General result 1. Let a decision theory D represent preferences over finite discrete lotteries L $L = \{x_j, p_j\}, x_{j'} \ge x_j$ for $j' > j$ with "utility functional"

(a.i) $$U(L) = \sum_{j\ge 1} v(x_j) \int_{P_{j+1}}^{P_j} df$$

where $P_j = \Pr(x : x \geq x_j\}$, $f(\cdot)$ is the transformation of decumulative probabilities, P_j, whereas $v(.)$ is the money transformation function. We use the normalization, $v(0) = 0$.

Suppose that statement P.1 as stated in Proposition 1 holds; that is

(a.ii) $\{y, p_i - \delta; x, 2\delta\}$ $\{y, p_i\}$, for all $i = 1, 2, \cdots, 2n-1$.

where $\delta = 1/2n$ and $\mathrm{p}_i = i/2n = i\delta$. Using notation $C \equiv v(y)/v(x)(>2)$ we show that according to theory D, statement P.1 implies that for all z, getting z for sure is preferred to getting $zK(C,n,n)$ or zero with even odds, for $K(.)$ as defined in section II.A. This suffices to derive that there are no D agents who satisfy both P.1 statement and Q.1 statement: preference for a binary lottery, $\{zK(C,n,n), 0.5\}$ against the sure amount of money z, for some z.

Proof. According to theory D, statement (a.ii) writes as

(a.1) $$v(x)\int_{(i-1)\delta}^{(i+1)\delta} df + v(y) f((i-1)\delta) \geq v(y) f(i\delta),\ i = 1, \ldots, 2n-1$$

Adding and subtracting $v(x) f(i\delta)$ and rearranging terms (a.1) becomes

(a.2) $$\int_{i\delta}^{(1+i)\delta} df \geq (C-1)\int_{(i-1)\delta}^{i\delta} df,\ i = 1, \ldots, 2n-1$$

Write inequality (a.2) for $i + k (= 1, \ldots, 2n)$ and apply it k times to get

$$\int_{(i+k)\delta}^{(i+k+1)\delta} df \geq (C-1)\int_{(i+k-1)\delta}^{(i+k)\delta} df \geq \ldots \geq (C-1)^k \int_{i\delta}^{(i+1)\delta} df$$

which generalizes as

(a.3) $$\int_{j\delta}^{(j+1)\delta} df \geq (C-1)^{j-i}\int_{(j-i)\delta}^{(j-i+1)\delta} df,\ \text{for all } j = i, \ldots, 2n.$$

To complete the proof it suffices to show that

(a.4) $$f(0.5) \leq \sum_{i=1}^{n}\left(\frac{1}{C-1}\right)^{i-1}\int_{0.5-\delta}^{0.5} df \quad \text{and} \quad 1 - f(0.5) \geq \sum_{j=1}^{n}(C-1)^j \int_{0.5-\delta}^{0.5} df$$

because two inequalities in (a.4) imply that $1 \geq f(0.5)\left[1 + \sum_{j=1}^{n}(C-1)^j / \sum_{i=1}^{n}(C-1)^{1-i}\right]$; hence, for any given z, the multiplication of both sides of the last inequality with $v(z)$ gives us the needed result: $v(z) \geq f(0.5)v(z)K(C,n,n) \geq f(0.5)v(zK(C,n,n))$. (For the last inequality, recall that $v(\cdot)$ is sub-additive.)

To show the first inequality of (a.4) verify that

$$f(0.5) = f(n\delta) = \sum_{i=1}^{n} \int_{(i-1)\delta}^{i\delta} df \le \sum_{i=1}^{n} \left(\frac{1}{C-1}\right)^{i-1} \int_{(n-1)\delta}^{n\delta} df = \sum_{i=1}^{n} \left(\frac{1}{C-1}\right)^{i-1} \int_{0.5-\delta}^{0.5} df$$

(The inequality follows from inequality (a.3)). For the second inequality of (a.4) verify that

$$1 - f(0.5) = f(2n\delta) - f(n\delta) = \sum_{j=n+1}^{2n} \int_{(j-1)\delta}^{j\delta} df \ge \sum_{j=1}^{n} (C-1)^{j} \int_{0.5-\delta}^{0.5} df$$

Proof of Proposition 1 (dual theory of expected utility).

Part (a) follows from the independence axiom. To show part (b) take any EU agent with $v(y) > 2v(x)$. This EU agent satisfies both P.1 and Q.1 with $z = y/G = y/K(y/x,n,n)$. Indeed, for $t > 2$ verify that $t < K(t,n,n)$; hence $v(zG) = v(y) > 2v(x) > 2v(y/G) = 2v(z)$.

Part (c) is a special case of the general result 1 for $v(z) = z$.

Proof of Corollary 1.1 (zero reference-dependent preferences).

It is a straightforward application of the general result 1 for $v(z) = \upsilon(z)$.

Proof of Corollary 1.2 (loss aversion with x reference-dependent preferences).

Let the reference point be x. We show that preference for three outcome lotteries implies that for any given positive z, the certainty equivalent of $\{zK_{R+1}, 0.5\}$ is not larger than z, where $R = -\upsilon(y-x)/\mu(-x) > 1$.

Proof. Suppose that a loss averse agent satisfies statement P.1 which states that

(a.iv) $\qquad \{y, p_i - \delta; x, 2\delta\} \qquad \{y, p_i\}$, for all $i = 1, 2, \cdots, 2n-1$.

If x is the reference point then statement (a.iv) implies

(a.5) $\qquad \mu(-x) f^{-}(1-(i+1)\delta) + \upsilon((y-x) f^{+}((i-1)\delta) \ge \mu(-x) f^{-}(1-i\delta) + \upsilon(y-x) f^{+}(i\delta),$

for all $i = 1, \dots, 2n-1$, which can be equivalently rewritten as

(a.6) $$\int_{i\delta}^{(i+1)\delta} df^{+} = \int_{1-(i+1)\delta}^{1-i\delta} df^{-} \ge \frac{\upsilon(y-x)}{-\mu(-x)} \int_{(i-1)\delta}^{i\delta} df^{+},\ i = 1, \dots, 2n-1.$$

(The first inequality follows from $f^{+}(p) = 1 - f^{-}(1-p)$.) Use notation R and apply the last inequality $j-i$ times to get

(a.7) $$\int_{j\delta}^{(j+1)\delta} df^{+} \ge R^{j-i} \int_{i\delta}^{(i+1)\delta} df^{+},\ \text{for all } j = i, \dots, 2n-1,$$

and then (as in the proof for the general result 1) verify that the following inequality is true

$$1 \geq f^{+}(0.5)\left[1+\sum_{j=1}^{n} R^{j} / \sum_{i=1}^{n} R^{1-i}\right]=K_{R+1} f^{+}(0.5).$$

Finally, for any given positive z, the last inequality and sub-additivity of $\upsilon(\cdot)$ imply

$\upsilon(z) \geq \upsilon(z) K_{R+1} f^{+}(0.5) \geq \upsilon(z K_{R+1}) f^{+}(0.5)$.

which requires that the certainty equivalent of $\left\{zK_{R+1}, 0.5\right\}$ be less than z .

Corollary 1.3 (not for all p statement)

Part a. Suppose that statement P.1 is satisfied only for $p_i < 1/2 + m/2n$, that is

(a.ii') $\{y, p_i - \delta; x, 2\delta\} \quad \{y, p_i\}$, for all $i = 1, 2, \cdots, n, n+1, ..., n+m$, where $m \in N$ $s.t.$ $m < n.$

The proof is the same as in the general result 1. The only changes are: (i) instead of (a.4) show

(a.4') $$f(0.5) \leq \sum_{i=1}^{n}\left(\frac{1}{C-1}\right)^{i-1} \int_{0.5-\delta}^{0.5} df \text{ and } 1-f(0.5) \geq \sum_{j=1}^{m}(C-1)^{j} \int_{0.5-\delta}^{0.5} df;$$

and that (ii) $1 \geq f(0.5)\left[1+\sum_{j=1}^{m}(C-1)^{j} / \sum_{i=1}^{n}(C-1)^{1-i}\right] = f(0.5)K(C,m,n)$.

(Note that a stronger Q.1 statement can be derived: (iii) $f(0.5 + m/2n) \geq f(0.5)K(C,m,n)$; hence $\{z, 0.5 + m/2n\} \succ \{zK(C,m,n), 0.5\}$ for all z.)

The second inequality of (a.4') follows from

$$1-f(0.5)=\sum_{j=n+1}^{2n} \int_{(j-1)\delta}^{j\delta} df = \sum_{j=n+1}^{n+m} \int_{(j-1)\delta}^{j\delta} df + \sum_{j=n+m+1}^{2n} \int_{(j-1)\delta}^{j\delta} df \geq \left(\sum_{j=1}^{m}(C-1)^{j-1}\right) \int_{0.5-\delta}^{0.5} df$$

Part b. Suppose that

(a.ii') $\{y, p_i - \delta; x, 2\delta\} \quad \{y, p_i\}$, for all $i = n, n+1, ..., 2n-1$.

Straightforwardly, it can be shown that statement (a.ii') implies $1 - f(0.5) + Gf(0.5) \geq Gf(0.75)$. Make only one change in the proof of the general result 1: instead of (a.4) write

(a.4') $$f(0.75)-f(0.5) \leq \sum_{i=1}^{n/2}\left(\frac{1}{C-1}\right)^{i-1} \int_{0.75-\delta}^{0.75} df \text{ and } 1-f(0.75) \geq \sum_{j=1}^{n/2}(C-1)^{j} \int_{0.75-\delta}^{0.75} df$$

Part c. Suppose that

(a.ii') $\{y, p_i - \delta; x, 2\delta\} \quad \{y, p_i\}$, for all $i = 1, ..., n$.

Statement (a.ii') implies $f(0.5) \geq Gf(0.25)$; in the proof of the general result 1, replace (a.4) with (a.4') $$f(0.25) \le \sum_{i=1}^{n/2} \left(\frac{1}{C-1} \right)^{i-1} \int_{0.25-\delta}^{0.25} df \text{ and } f(0.5) - f(0.25) \ge \sum_{j=1}^{n/2} (C-1)^j \int_{0.25-\delta}^{0.25} df$$

A.3. Proof of Proposition 2 and Corollary 2

General result 2. Let a decision theory D with "utility functional" U as in statement (a.i) be given. We assume here that v is (weakly) concave and differentiable (the proof extends straightforwardly to non-differentiable weakly concave functions; see also Rabin, 2000.). Denote Denote $a = \ell$ and $b = g + \ell$. Statement P.2 in this notation is

(a.iv) $x + a \qquad \{x + b, p; x\}$ for all integers $x \in (m, m + N(l + g))$, $m > 0$.

For a general p, condition $g > \ell$ generalizes to

(a.v) $bf(p) > a$.

We show that (a.iv) and (a.v) imply $z = m + K(g + \ell) \succ \{z + J(g + \ell), p; m\}$, where generalized J is defined in (*) in Proposition 2 by replacing ℓ / g with $q = (1 - f(p))\ell / f(p)g (= r(f(p))$, see Section III.B for the specification of r(.)); that is (*) $J = \lceil N - K + 1 + A(q, K)q^{-N} \rceil$.

Proof. Condition (a.iv) implies

(a.6) $v(x + a) \ge (1 - f(p))v(x) + f(p)v(x + b)$, for all $x \in (m, m + Nb)$.

The proof consists of two steps: First, we show that (a.6) and concavity of $v(\cdot)$ imply that for all $y \in (m, m + Nb)$

(a.7) $v'(y + jb) \le q^j v'(y)$, for all $j \in \Psi_y = \{j \in N \mid y + (j-1)b \in (m, m + Nb)\}$;

Then, we show that for any given $z = m + Kb, K < N$

(a.8) $v(z) = v(m + Kb) \ge f(p)v(m + Kb + Jb) + (1 - f(p))v(m) = f(p)v(z + Jb) + (1 - f(p))v(m)$

which completes the proof.

To derive (a.7): write $v(x + a) = f(p)v(x + a) + (1 - f(p))v(x + a)$, rewrite (a.6) with $x = y$, and rearrange terms to get

(a.9) $(1 - f(p))\left[v(y + a) - v(y)\right] \ge f(p)\left[v(y + b) - v(y + a)\right]$, $\forall y \in (m, m + Nb)$.

Statement (a.9), inequalities $\left[v(y + b) - v(y + a)\right] / (b - a) \ge v'(y + b)$, $\left[v(y + a) - v(y)\right] / a \le v'(y)$, (both following from the weak concavity of $v(\cdot)$) imply

(a.10) $v'(y + b) \le qv'(y)$, $\forall y \in (m, m + Nb)$.

Finally, to get statement (a.7) simply iterate the last inequality, (a.10) j times.

To show statement (a.8), first verify that

(a.11) $$v(z)-v(z-bK)=\sum_{k=0}^{K-1}\left[v(z-kb)-v(z-(k+1)b)\right]\geq b\sum_{k=0}^{K-1}v'(z-kb)\geq bv'(z)\sum_{k=0}^{K-1}(1/q^{k})$$

Next, note that $K>2-\ln 2/\ln q$ implies $A(r,K)>0$ and hence $J+K>N$. Weak concavity of $v(\cdot)$ and statement (a.7) imply

(a.12) $$\begin{aligned}&v(z+Jb)-v(z)=\sum_{j=0}^{J-1}\left[v(z+(j+1)b)-v(z+jb)\right]\\&\leq b\left[(J-N+K)v'\big(z+(N-K)b\big)+\sum_{j=0}^{N-K-1}v'(z+jb)\right]\leq bv'(z)\left[q^{N-K}(J-N+K)+\sum_{j=0}^{N-K-1}q^{j}\right]\end{aligned}$$

Statements (a.11) and (a.12) imply that a sufficient condition for (a.8) is

(a.13) $$(1-f(p))\sum_{k=0}^{K-1}(1/q^{k})\geq f(p)\left[q^{N-K}(J-N+K)+\sum_{j=0}^{N-K-1}q^{j}\right]$$

Substitute $(1-q)\sum_{j=0}^{N-K-1}q^{j}=1-q^{N-K}$, and $(1-q)\sum_{k=0}^{K-1}q^{-k}=q^{1-K}-q$ in (a.13) to get

(a.14) $$J<N-K+\left(\frac{1-f(p)}{f(p)}(q^{1-K}-q)-(1-q^{N-K})\right)\frac{1}{(1-q)q^{N-K}}$$

The last inequality is true because

$$J=\left\lceil N-K+1+A(q,K)q^{-N}\right\rceil\leq N-K+\frac{1}{1-q}(1+A(q,K)q^{-N})=N-K+\frac{1}{1-q}(1+(q^{2}-q^{2+K}-q^{K})q^{-N})$$

$$<N-K+\frac{1}{1-q}(1+(\frac{g}{\ell}q(q-q^{1+K})-q^{K})q^{-N})=N-K+\frac{1}{1-q}\left[1+(\frac{1-f(p)}{f(p)}(q-q^{1+K})-q^{K})q^{-N}\right]$$

where the first equality is true by the construction of *J*, the first inequality holds for $q\in(0,1)$, the second equality follows from the definition of $A(q,K)$, the second inequality follows from $g>\ell$, and the last equality follows from notation $q=(1-f)\ell/fg$.

Proof of Proposition 2 (expected utility theory).

Part a. Use linearity in payoffs of the DTEU functional for a straightforward derivation.

Part b. It can be verified that any DTEU agent with $f(0.5)\in I=\big(K/(N+1),\ell/(g+\ell)\big)$ satisfies both P.2 and Q.2. ($I\neq\phi$ follows from $N>(1+g/\ell)N_{*}(\ell/g)$.)

Part c. It is an application of the general result 2 for p = 0.5, $f(p)=p$, $q=\ell/g$ and $v(z)=u(z)$.

Corollary 2 (cumulative prospect theory and rank dependent utility theory).

It is an application of the general result 2 for p = 0.5, $f(p)=h(p)$ and $v(z)=\upsilon(z)$.