Is There A Plausible Theory for Risky Decisions?

James C. Cox[1], Vjollca Sadiraj[1], Bodo Vogt[2], and Utteeyo Dasgupta[3]

1. **Georgia State University**

2. **University of Magdeburg**

3. **Franklin and Marshall College**

Is There A Plausible Theory for Risky Decisions?

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

A large literature is concerned with analysis and empirical application of theories of decision making for environments with risky outcomes. Expected value theory has been known for centuries to be subject to critique by St. Petersburg paradox arguments. More recently, theories of risk aversion have been critiqued with calibration arguments applied to concave payoff transformations. This paper extends the calibration critique to decision theories that represent risk aversion solely with transformation of probabilities. Testable calibration propositions are derived that apply to four representative decision theories: expected utility theory, cumulative prospect theory, rank-dependent expected utility theory, and dual expected utility theory. Heretofore, calibration critiques of theories of risk aversion have been based solely on thought experiments. This paper reports real experiments that provide data on the relevance of the calibration critiques to evaluating the plausibility of theories of risk aversion. The paper also discusses implications of the data for (original) prospect theory with editing of reference payoffs and for the new dual-self model of impulse control. In addition, the paper reports an experiment with a truncated St. Petersburg bet that adds to data inconsistent with risk neutrality.

Keywords: Risk, Calibration, Decision Theory, Game Theory, Experiments

1. Introduction

A large literature in economics, finance, game theory, management science, psychology, and related disciplines is concerned with analysis and empirical application of theories of decision making for environments with risky outcomes. Together, results from old and recent thought experiments question the existence of a plausible theory of risky decisions. We report real experiments that address this question.

The first theory of risk-taking decisions, expected value maximization, was challenged long ago by the St. Petersburg paradox (Bernoulli, 1738). This traditional critique of the plausibility of expected value theory is based on hypothetical experiments in which people report that they would not be willing to pay more than a moderate amount of money to play a St. Petersburg game with infinite expected value. A traditional defense of the theory is based on the observation that no agent can credibly offer the St. Petersburg game for another to play in a real-payoff experiment – because it could result in a payout obligation exceeding any agent's wealth – and therefore that this challenge to expected value theory has no bite. Instead of participating in this traditional debate, we present a feasible St. Petersburg game and report data from its implementation.

Whether or not decisions under risk are consistent with risk neutrality is important for many applications of theory, including typical applications of game theory in which the payoffs are (often implicitly) assumed to be risk neutral utilities (Roth, 1995, pgs. 40-49). Of course, abstract (as distinct from applied) game theory does not require risk neutrality because the abstract payoffs *are* utilities. But if expected value theory is implausible then one needs an alternative theory of utility for game theory. This can explain why von Neumann and Morgenstern (1947) developed both a theory of utility and a theory of play for strategic games. From a contemporary perspective one understands that other theories of utility, different from expected utility theory, can be incorporated into game theory. But if all known theories of utility have implausibility problems then this is a general problem for game theory not only a problem for (game-against-nature) applications of theory to explain individual agents' decisions in risky environments.

Oral tradition in economics historically held that decreasing marginal utility of money could not explain non-risk-neutral behavior for small-stakes risky money payoffs. Recently, criticism of decreasing marginal utility of money as an explanation of risk aversion has been formalized in the concavity calibration literature sparked by Rabin (2000). Papers have explored the implications of calibration of concave payoff transformations for the expected utility of terminal wealth model (Rabin, 2000) and for expected utility theory, rank-dependent expected utility theory, and cumulative prospect theory (Neilson, 2001; Cox and Sadiraj, 2006).

This paper demonstrates that the problem of implausible implications from theories of decision making under risk is more generic, and hence more fundamental, than implausible (implications of) decreasing marginal utility of money. To make this point, we consider a model with constant marginal utility of money that explains risk aversion solely with decreasing marginal significance of probability. We extend the calibration literature to include the implications of convex transformations of decumulative probabilities used to model risk aversion in Yaari's (1987) dual theory of expected utility. Convex transformation of decumulative probabilities implies concave transformation of cumulative probabilities, which in the case of binary lotteries overweights the low payoff and underweights the high payoff (which can be characterized as "pessimism").

In order to explicate the general nature of the implausibility problem, we also consider rank dependent expected utility theory (Quiggin, 1982, 1993) and cumulative prospect theory (Tversky and Kahneman, 1992). We present calibration propositions that apply to four representative decision theories that explain risk aversion: (a) solely with concave transformations of payoffs (expected utility theory) or (b) solely with convex transformations of decumulative probabilities (dual theory) or (c) with both payoff and probability transformations (rank dependent expected utility theory and cumulative prospect theory).

Risk aversion calibrations are based on assumptions about patterns of risk aversion. Previous conclusions about the implications of concavity calibration for implausibility of the expected utility of terminal wealth model have been based on thought experiments with empirical validity of the assumption that a given bet will be rejected over a large range of initial wealth levels (Rabin, 2000). Cox and Sadiraj (2006) explore alternative assumed patterns of risk aversion that have concavity-calibration implications for cumulative prospect theory as well as three expected utility models (expected utility of terminal wealth, expected utility of income, and expected utility of initial wealth and income) but do not report any data. This paper reports real experiments designed to shed light on the empirical validity of the risk aversion assumptions underlying Propositions 1 and 2 below, and thereby on the relevance of these calibrations to evaluating the empirical plausibility of theories of risk aversion.

The risk aversion assumption underlying Proposition 1 provides the basis for an experimental design that can implement within-subjects tests for implausible implications of calibration for models with decreasing marginal utility of income. A different risk aversion assumption underlying Proposition 2 provides guidance for design of a within-subjects test for implausible implications of calibration for models with decreasing marginal significance of probability.

We report data from three risk-taking experiments that, according to the calibration propositions, have implications of implausible risk aversion in the large for expected utility theory, cumulative prospect theory, rank-dependent expected utility theory, and dual expected utility theory. We also explain implications of the data for the original version of prospect theory (Kahneman and Tversky, 1979) in which "editing" can arguably be used to immunize the theory to the implications of

concavity calibration but not to other implications inherent in our experimental designs. Finally, we ask whether the dual-self model (Fudenberg and Levine, 2006) can rationalize data from our experiments.

2. Risk Neutrality in Laboratory Experiments

Despite the existence of a large amount of real-payoff experiment data inconsistent with risk neutrality, the question has not been completely resolved. Much of this literature involves testing of compound hypotheses involving risk neutrality and additional context-specific assumptions about behavior (such as Nash equilibrium strategies). We report an experiment with finite St. Petersburg bets that supports tests for risk neutrality as a simple hypothesis.

2.1 An Experiment with Truncated St. Petersburg Bets

The experiment was designed as follows. Subjects were offered the opportunity to decide whether to pay their own money to play nine truncated St. Petersburg bets. One of their decisions (for one of the bets) was randomly selected for real money payoffs. Bet N had a maximum of N coin tosses and paid 2^n euros if the first head occurred on toss number n, for $n = 1, 2, ... N$, and paid nothing if no head occurred. Bets were offered for $N = 1, 2, …, 9$. Of course, the expected payoff from playing bet N was N euros. The price offered to a subject for playing bet N was 25 euro cents lower than N euros. An expected value maximizer would accept these bets.

The experiment was run at the University of Magdeburg in February 2007. An English version of the instructions is available on an author's homepage.[1] Thirty subjects participated in this experiment. As reported in Table 1: (a) 127 out of 220 (or 47%) of the subjects' choices are inconsistent with risk neutrality; and (b) 26 out of 30 (or 87%) of the subjects made at least one choice inconsistent with risk neutrality.[2] Therefore, this experiment reveals substantial aversion to risks involving the opportunity to pay N - 0.25 euros to play a truncated St. Petersburg bet with expected value N euros and highest possible payoff of 2^N euros, for $N = 1$ 2, …,9.

2.2 Existing Data on Small-Stakes Risk Aversion

An experiment by Holt and Laury (2002) generated data that support tests for risk neutrality. In their small-stakes treatment, they asked subjects to make choices between two risky alternatives in each of 10 rows. In row i, for $i = 1, 2, ..., 10$: option A paid \$2.00 with probability $i/10$ and paid \$1.60 with probability $1 - i/10$; option B paid \$3.85 with probability $i/10$ and paid \$0.10 with probability $1 - i/10$. One row was randomly selected for money payoff. An expected value maximizer would choose option A in rows 1-4 and option B in rows 5-10. A rational risk averse agent would switch once from choosing option A to choosing option B in some row (weakly) between rows 5 and 10.

Holt and Laury reported that a large proportion of their subjects chose option A in row 5, as well as rows 1-4, and switched to option B in row 6 or a higher numbered row. Such behavior is consistent with risk aversion but inconsistent with risk neutrality for small-stakes. Holt and Laury (2002) also reported data generally inconsistent with risk neutrality from a similar risk treatment in which payoffs were scaled up by a factor of 10. Harrison (2005) examined possible treatment order effects in the Holt and Laury (2002) experiment. Harrison (2005) and Holt and Laury (2005) report data from experimental designs intended to mitigate possible treatment order effects. Data from all of the experiments support the conclusion that a high proportion of subjects are risk averse even for small-stakes lotteries. Two of our calibration experiments, explained below, add more data inconsistent with risk neutrality.

3. Calibration of Payoff Transformations

Data inconsistent with risk neutrality provide support for the need for theories of risk averse decision making. Search for a unified theory of risky decision making leads one to try to develop a theory with empirical validity that can be applied in both experimental and naturally-occurring risky environments. This quest for a unified decision theory of risky decisions introduces calibration issues.

Choice according to expected value maximization corresponds to maximizing a utility functional that is linear in both payoffs and probabilities. Risk aversion is incorporated into decision theories by "utility functionals" characterized by nonlinear transformation of payoffs and/or nonlinear transformation of probabilities.[3] We consider four representative examples. As is well known, expected utility theory incorporates risk aversion through strictly concave transformation of payoffs.

In contrast, dual expected utility theory (Yaari, 1987) incorporates risk aversion solely through strictly convex transformation of decumulative probabilities while rank dependent expected utility theory (Quiggin, 1982, 1993) and cumulative prospect theory (Tversky and Kahneman, 1992) use nonlinear transformations of both payoffs and probabilities.

Rabin's (2000) assumed patterns of risk aversion have concavity-calibration implications that apply to the expected utility of terminal wealth model. His concavity calibration is based on the assumption that an agent will reject a 50/50 bet of losing $\ell\,(>0)$ or gaining $g\,(>\ell)$ at all initial wealth levels in a (finite or infinite) interval. Cox and Sadiraj's (2006) assumed patterns of risk aversion have concavity-calibration implications for cumulative prospect theory and three expected utility models (of terminal wealth, income, and initial wealth and income), but only for infinite income intervals. Their concavity calibration proposition is based on the assumption that an agent will reject a 50/50 bet of gaining $x-\ell\ or\ x+g,\,(g>\ell>0)$ in favor of receiving x for sure, for all amounts of income x greater than ℓ. In order to test several representative theories of risk averse decision making, one needs a concavity calibration that applies to *finite* (income or wealth) intervals and to "utility functionals" that are or are not linear in probabilities. We here present just such a concavity calibration proposition. Before stating the propositions, we first introduce some definitions.

Let $\{x, p; y\}$ denote a binary lottery that yields payoff x with probability p and payoff y with probability $1-p$. Now consider a decision theory D that represents a preference ordering of binary lotteries $\{x, p; y\}$, for x < y, with a "utility functional" F_D given by

(1) $\qquad F_D(x, p; y) = h(p)\varphi(x) + (1 - h(p))\varphi(y)$.

The lottery $\gamma = \{a, p; b\}$ is preferred to lottery $\delta = \{c, \rho; d\}$ if and only if $F_D(a, p; b) > F_D(c, \rho; d)$. Using (1) this lottery preference can be explicitly written as

(2) $\qquad \gamma \succ \delta$ iff $h(p)\varphi(a) + (1 - h(p))\varphi(b) > h(\rho)\varphi(c) + (1 - h(\rho))\varphi(d)$.

The function h is a probability transformation (or weighting) function and the function φ is a payoff transformation function (e.g., a Bernoulli utility function or prospect theory value function). For the special case of expected utility theory, the probability transformation function is the identity map $h(p) = p$ as a consequence of the independence axiom. For the special case of dual expected utility

theory, the payoff transformation function is the identity map $\varphi(z) = z$ as a consequence of the dual independence axiom. For cumulative prospect theory or rank dependent expected utility theory or dual expected utility theory, $h(p)$ is a probability weighting function.[4]

The lottery $\{x, p; y\}$ is said to be $z_D -$ favorable if it satisfies

(3) $h(p)x + (1 - h(p))y > z$.

In the special case of expected utility theory, a lottery is $z_D -$ favorable if and only if its expected value is larger than z.

Define the variable q with the statement

(4) $q = \dfrac{h(p)}{1 - h(p)} \dfrac{a}{b - a}$.

For $0 < a < b$ given such that $\{0, p; b\}$ is an $a_D -$ favorable lottery, statements (3) and (4) imply that $q < 1$. In the following proposition the notation $\langle x \rangle$ denotes the largest integer smaller than x.

Proposition 1. *Given an $a_D -$favorable lottery $\{0, p; b\}$, suppose that an agent prefers the certain amount of money $x + a$ to lottery $\{x, p; x + b\}$, for all $x \in [m, M]$. Then according to decision theory D, for any given $Z \in (m, M)$ such that $q^{\langle (Z-m)/b \rangle - 1} < h(p)$, the agent prefers the certain amount of money Z to the lottery $\{m, p; G\}$ for all G such that*

(*) $G < M + b(2q - 1)/(1 - q) + Cq^{-\langle (M-m)/b \rangle}$,

where q is defined in statement (4) and $C = \left((b/a - 1)q^2 (1 - q^{\langle (Z-m)/b \rangle}) - q^{\langle (Z-m)/b \rangle} \right) b/(1 - q)$.

Proof: See appendix A.

Note that for any given *m* and *Z*, the third term on the right hand side of inequality (*) increases geometrically in M because $q < 1$ (which follows from $\{0, p; b\}$ being an a_D -favorable lottery). This implies that for any amount of gain *G*, as big as one chooses, there exists a large enough interval in which preference for $x + a$ over a risky lottery $\{x, 0.5; x + b\}$, for all *x* from the interval $[m, M]$ implies a preference for *Z* for sure to the risky lottery $\{m, 0.5; G\}$. We use inequality (*) in Proposition 1 to construct Table 2. The figures in Table 2 can be used to explain some implications of

concavity calibration with a_D – favorable lotteries for expected utility theory, rank dependent expected utility theory and cumulative prospect theory, as follows.

3.1 Implications for the Expected Utility of Terminal Wealth Model.

Consider the lottery first discussed by Rabin (2000). Assume that an agent rejects the lottery in which he would lose 100 or gain 110, with equal probability, for all values of initial wealth ω in the interval [1000, *M*], where values of *M* are given in the first column of Table 2. Then the expected utility of terminal wealth model predicts that the agent prefers the certain income 3000 to a risky lottery $\{1000, 0.5; G\}$ where the values of *G* are given in First EU Calibration (or second) column of Table 2. As a specific illustrative example, consider the *M* = 30,000 (or fifth) row entry in column 2. This entry in Table 2 informs us that an expected utility of terminal wealth maximizing agent who rejects the lottery $\{-100, 0.5; +110\}$ at all initial wealth levels $\omega \in (1000, 30000)$, would also reject the 50/50 lottery in which he would lose 1000 or gain 110 million at all $\omega \geq 3,000$. This level of implied risk aversion is implausible.

3.2 Implications for Other Expected Utility Models

Proposition 1 implies that the figures in the First EU Calibration (or second) column of Table 2 also apply to the expected utility of income model and the expected utility of initial wealth and income model (discussed in Cox and Sadiraj, 2006) *if one assumes* that an agent prefers the certain amount of income $x+100$ to the lottery $\{x, 0.5; x+210\}$, for all $x \in [900, M]$, where values of *M* are given in the first column of Table 2. Then *all three* expected utility (of terminal wealth, income, and initial wealth and income) models predict that the agent prefers receiving the amount of income 3000 for sure to a risky lottery $\{900, 0.5; G\}$, where the values of *G* are given in the second column of Table 2. For example, if $[m, M] = [900, 50000]$ then $G = 0.1 \times 10^{13}$ for all three expected utility models, which is implausible risk aversion. Concavity calibration with this lottery has no implication for cumulative prospect theory if one uses the probability weighting function estimated by Tversky and Kahneman

(1992, p. 312). They report that the probability weighting function for gains is such that $W^+(0.5) = 0.42$. Setting $h(0.5) = 0.58$ $(= 1 - W^+(0.5))$ reveals that the lottery $\{0, 0.5; 210\}$ is not 100_D-favorable for prospect theory and therefore Proposition 1 does not apply since statement (4) is not satisfied. However, the other columns in Table 2 do report concavity calibration implications for cumulative prospect theory, rank dependent expected utility theory, and all three of the expected utility models mentioned above.

3.3 Implications for Cumulative Prospect Theory and Rank Dependent Expected Utility Theory

Proposition 1 has implications for cumulative prospect theory and rank dependent expected utility theory as well as expected utility theory, as can be seen from interpreting the probability weighting function $h(p)$ in statements (1) - (4) above. For expected utility theory $h(p) = p$, *for all* $p \in [0,1]$, as a consequence of the independence axiom. For rank dependent expected utility theory, our $h(p)$ is equivalent to Quiggin's (1993, p.52) $q(p)$ *for binary lotteries*. For cumulative prospect theory, our $h(p)$ is equivalent to Tversky and Kahneman's (1992, p.300) $1 - W^+(p)$ *for binary lotteries*.

Assume that an agent prefers the degenerate lottery $\{x + 100, 1; 0\}$ to the lottery $\{x, 0.5; x + 250\}$ for all $x \in [900, 30000]$. Lottery $\{0, 0.5; 250\}$ is 100_D − favorable for expected utility theory ($h(0.5) = 0.50$), cumulative prospect theory ($h(0.5) < 0.6$), and rank dependent expected utility theory (for $h(0.5) > 0.40$) by statement (3). According to the third column entry in the $M = 30,000$ row of Table 2, expected utility theory predicts that such an agent will reject a 50/50 lottery with positive outcomes 900 or 0.12×10^{24} in favor of getting an amount of income 3,000 for sure. As shown in the fourth column of Table 2, rank dependent expected utility theory (with $h(0.5) = 0.42$) and cumulative prospect theory (with $h(0.5) = 0.58$) predict that such an agent will reject a 50/50 lottery with positive outcomes 900 or 0.46×10^7 in favor of an amount of income 3,000 for sure, which again is implausible risk aversion.

As a final illustrative example, assume that an agent prefers the degenerate lottery $\{x + 20, 1; 0\}$ to the lottery $\{x, 0.5; x + 50\}$, for all $x \in [900, 6000]$. Lottery $\{0, 0.5; 50\}$ is

$20_D -$ favorable for expected utility theory, rank-dependent expected utility theory, and cumulative prospect theory by statement (3). According to the fifth column entry in the $M = 6{,}000$ row of Table 2, expected utility theory predicts that an agent who rejects these lotteries will also reject a very large-stakes lottery with 0.5 probabilities of gaining 1,000 or gaining 0.4×10^{20} in favor of getting an amount of income 3,000 for sure. As shown in the sixth column of Table 2, cumulative prospect theory and rank dependent expected utility theory predict than an agent who rejects these same lotteries will also reject a very large-stakes lottery with 0.5 probabilities of gaining 1,000 or gaining 0.29×10^{7} in favor of an amount of income 3,000 for sure.

4. Calibration of Probability Transformations

The preceding section explains the implausible implications of arguably-plausible patterns of risk aversion for decision theories with "utility functionals" with concave transformations of payoffs. The discussion is based on section 3 applications of Proposition 1. This proposition has no implications for one prominent theory of risk aversion, Yaari's (1987) dual theory of expected utility, because it has constant marginal utility of income. The dual theory has a "utility functional" that is always linear in payoffs but it is linear in probabilities if and only if the agent is risk neutral. Risk aversion is represented in the dual theory by convex transformations of decumulative probabilities. Proposition 2 presents a calibration that applies to this theory.

Let $\{x_1, p_1; x_2, p_2; \cdots; x_{k-1}, p_{k-1}; x_k\}$, $\sum_{i=1}^{k-1} p_i \in [0,1]$ denote the lottery that gives prize x_i with probability p_i, for $i = 1, 2, \cdots, k-1$, and prize x_k with probability $1 - \sum_{i=1}^{k-1} p_i$.

Proposition 2. *For any given number* n, *define* $\delta = 1/2n$. *Suppose that for some* $c > 2$ *an agent prefers lottery* $R(i) = \{cx, (i-1)\delta; x, 2\delta; 0\}$ *to lottery* $S(i) = \{cx, i\delta; 0\}$ *for all* $i = 1, 2, \cdots, 2n-1$. *Then dual expected utility theory predicts that the agent prefers getting a positive* z *for sure to the lottery* $\{Gz, 0.5; 0\}$ *where* $G = 1 + \sum_{i=1}^{n} (c-1)^i \Big/ \sum_{j=1}^{n} (c-1)^{1-j}$.

Proposition 2 states the following. Suppose that an agent prefers a lottery with outcomes 0, x, and $c \times x$ with probabilities $1 - (i+1)\delta$, 2δ, and $(i-1)\delta$, respectively, to a $(1-i\delta)/i\delta$ lottery with outcomes 0 and $c \times x$, for all $i = 1, 2, \ldots, 2n-1$. Then according to dual theory the agent must prefer getting $z \ (> 0)$ for sure to a 50/50 lottery with outcomes 0 and Gz, where G is the entry in row c and column n of Table 3. For example, let $\delta = 0.05$ (i.e., $n = 10$) and $c = 4$. Then if an agent rejects lottery $\{40, i/20; 0\}$ in favor of lottery $\{40, (i-1)/20; 10, 0.1; 0\}$ for all $i = 1, \ldots, 19$, then the dual theory predicts that the agent prefers 100 for sure to 50/50 lottery with prizes 5.9 million or 0. For another example, let $\delta = 0.1$ (i.e., $n = 5$) and $c = 4$. Then if an agent rejects lottery $\{40, i/10; 0\}$ in favor of lottery $\{40, (i-1)/10; 10, 0.2; 0\}$ for all $i = 1, \ldots, 9$, then the dual theory predicts that the agent prefers 100 for sure to 50/50 lottery with prizes 24,400 and 0.

5. Experimental Design Issues

Some design problems are inherent in calibration experiments. The issues differ with the assumption underlying a calibration and the type of theory of risk aversion the calibration applies to.

5.1 Alternative Payoff Calibrations and Across-subjects vs. Within-subjects Designs

The analysis by Rabin (2000) is based on the assumption that an agent will reject a 50/50 bet with loss $\ell > 0$ or gain $g > \ell$ at all initial wealth levels w in a large (finite or infinite) interval. Rabin conducted a thought experiment on the empirical validity of his assumption that many readers found convincing. An attempt to conduct a real experiment on the assumption's empirical validity would encounter considerable difficulties. Unless an experiment was conducted over a many-year time horizon it would have to use an across-subjects design because, obviously, an individual's initial wealth is (approximately) constant during the short time frame of most experiments. Conducting either a many-year, within-subjects experiment or an ordinary (short-time-frame), across-subjects experiment would produce data containing confounding differences in demographic determinants of risk attitude in addition to differences in wealth. In addition the data would have no concavity-

calibration implications for decision theories in which income rather than terminal wealth is the postulated argument of the theory's utility functional.

Our Proposition 1 is based on the assumption that an agent will reject a 50/50 bet with gain x or gain $x+a$ in favor of a certain income in amount $x+b$, with $a > b > 0$, for all x in an interval $0 < m \le x \le M$. Calibrating the implications of our assumption follows the same logic as calibrating the implications of Rabin's assumption; in that sense the two assumptions are "mathematically equivalent." But the two assumptions have quite different empirical implications. An experiment on the empirical validity of our assumption can be conducted within a short time frame, using a within-subjects design, by varying income level x as a treatment parameter. This approach avoids confounds from changing demographic determinants of risk attitudes associated with passage of time (for a within-subjects design) of different personal characteristics (for an across-subjects design). In addition, with a suitable choice of the parameters a and b, the data have concavity-calibration implications for the expected utility of terminal wealth model and for other models, in which income is the argument of utility functionals, such as cumulative prospect theory and the expected utility of income model (Cox and Sadiraj, 2006).

5.2 Affordability vs. Credibility with Payoff Calibration Experiments

Table 2 illustrates the relationship between the size of the interval [m, M], in column (1), used in the assumption underlying a concavity calibration, and the size of the high gain G in the result reported in columns 2 - 6. If it were considered credible to exclusively run hypothetical payoff experiments, then there would be no difficult experimental design tradeoffs; one could choose [m, M] = [900, 70000], run experiments with all of the lotteries used in the concavity calibration reported in Table 2, and use the entries in the rows of the table to draw conclusions about plausible or implausible models of risk preferences. Economists are skeptical that data from hypothetical payoff experiments on risk-taking behavior are credible. But calibration experiments with money payoffs involve some difficult tradeoffs between what is affordable and what is credible, as we shall next explain.

As an example, suppose one were to consider implementing an experiment in which subjects were asked to choose between a certain amount of money $\$x + \$100$ for sure and the binary lottery

$\{\$x, 0.5; \$x + \$210\}$ for all x varying uniformly between $m = \$900$ and $M = \$350,000$. Suppose the subject always chooses the certain amount $\$x + \$100$ and that one decision is selected randomly 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? The implications of concavity calibration are 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 experimental design becomes clear from close scrutiny of Proposition 1: the unit of measure for m and M is the same as that for the amounts at risk, a and b in the binary lottery (see statement (*) in Proposition 1). If the unit of measure for m and M is $1/10,000 then the unit of measure for a and b is the same (or else the calibration doesn't apply); in that case the binary lottery would become $\{x, 0.5; x + \$0.021\}$, which involves only a 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 calibration in Proposition 1 to lead to the implication of implausible risk aversion in the large *if* the assumption underlying the calibration has empirical validity. There is no perfect solution to the problem. We implemented alternative imperfect solutions, as we explain in sections 6 and 8.

5.3 Saliency vs. Power with Probability Calibration Experiments

Table 3 illustrates the relationship between the scale of payoffs in the lotteries (x), the ratio of high and middle payoffs in the risky lottery (c), the difference between probabilities of high and low payoffs in adjacent terms in the calibration (determined by z in $\frac{i}{10^z} - \frac{i-1}{10^z}$).

The design problem for convexity calibration experiments with money payoffs is inherent in the need to have a fine enough partition of the [0,1] interval for the calibration in Proposition 2 to lead to the implication of implausible risk aversion in the large *if* the risk aversion assumption underlying

the calibration has empirical validity. For example if the length of each subinterval is $1/10^3$ then adjacent probabilities differ by 0.001 and the subjects' decision task is to make 1,000 choices; in such a case, the subjects would not be sensitive to the probability differences and the payoffs would arguably not be salient because of the huge number of choices needing to be made. In contrast, if the length of each subinterval is $1/10$ then the adjacent probabilities differ by 0.1 and the subjects' decision task is to make 10 choices. The calibration implications of the $1/10$ length of the subinterval are less spectacular but the resulting experimental design can be implemented.

6. Calcutta Experiment for Payoff Transformation Theories

An experiment with money payoffs was conducted at the Indian Statistical Institute in Calcutta during the summer of 2004. The subjects were resident students at the institute. Two sessions were run, each with 15 resident student subjects. E-mail announcements were sent a week in advance to recruit for student subjects. Subjects replied with their availability. Students who confirmed their availability on either of the two dates were recruited on a first-reply, first-served basis. Since the Indian Statistical Institute does not have an experimental laboratory, the sessions were run in a big lecture room with a capacity of about 80 people. This allowed for a convenient separation distance between the subjects. In this experiment, all payoffs were denominated in Indian rupees.

6.1 Experimental Design

In each experiment session a subject was asked to perform two tasks. For the first task, subjects were asked to make choices on six individual response forms, between a certain amount of money, x rupees + 20 rupees and a binary lottery, {x rupees, 0.5; x rupees + 50 rupees} for values of x from the set {100, 1K, 2K, 4K, 5K, 6K}, where $K = 1,000$. On each response form, subjects were asked to choose among option A (the risky lottery), option B (the certain amount of money), and option I (indifference). The alternatives given to the subjects are presented in Table 4. The second task was completion of a questionnaire including questions about amounts and sources of income. Appendix B.1 contains detailed information on the protocol of this experiment.

6.2 Economic Significance of the Certain Incomes and Lottery Risks

The exchange rate between the Indian rupee and the U.S. dollar at the time the Calcutta experiment was run was about 42 to 1. This exchange rate can be used to convert the rupee payoffs discussed above into dollars. Doing that would not provide very relevant information for judging the economic significance to the subjects of the certain payoffs and risks involved in the Calcutta experiment because there are good reasons for predicting that none of the subjects would convert their payoffs into dollars and travel to the United States to spend them. Better information on the economic significance of the payoffs to subjects is provided by comparing the rupee payoffs in the experiment to rupee-denominated monthly stipends of the student subjects and rupee-denominated prices of commodities available for purchase by students residing in Calcutta.

The student subjects' incomes were in the form of scholarships that paid 1,200-1,500 rupees per month for their expenses in addition to the standard tuition waivers that each received. This means that the highest certain payoff used in the experiment (6,000 rupees) was equal to four or five months' salary for the subjects. The daily rate of pay for the students was $40 - 50$ rupees. This means that the size of the risk involved in the lotteries (the difference between the high and low payoffs) was greater than or equal to a full day's pay.

A sample of commodity prices in Calcutta at the time of the experiment (summer 2004) is reported in Table 5. Prices of food items are reported in number of rupees per kilogram. 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. This implies that there are about 15 servings in a kilogram of these food items. As reported in Table 5, for example, we observed prices for poultry of $45 - 50$ rupees per kilogram. This implies that the size of the risk involved in the lotteries (50 rupees) was equivalent to 15 servings of poultry. The price of a moderate quality restaurant meal was $15 - 35$ rupees per person. This implies that 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.

6.3 Implications of the Data for Expected Utility Theory, Rank-Dependent Expected Utility Theory, and Cumulative Prospect Theory

There were in total 30 subjects in this experiment. Nine subjects never rejected a risky lottery whereas five subjects rejected the risky lottery only in the first decision task. There were eight subjects who revealed an interval of risk aversion with length at least $3.9K$. The expected utility, rank-dependent expected utility, and cumulative prospect theory calibration implications for these individuals are reported in Table 6. Figures reported in the third and fourth column of the table reveal implausible risk aversion calibration implications for these eight individuals. We conclude that 27% of the subjects satisfy the risk aversion assumption in intervals large enough to generate implausible risk aversion in the large for expected utility theory, rank-dependent expected utility theory, and cumulative prospect theory. Therefore, none of these theories provides a plausible theory of risk-taking decisions for 27% of the subjects in this experiment.

6.4 Implications of the Data for Non-cumulative Prospect Theory with Editing

In their later writings, on cumulative prospect theory, Kahneman and Tversky dropped some of the elements of the original version of the theory (Kahneman and Tversky, 1979). One element of the original version of prospect theory, known as "editing," can be described as follows. In comparing two prospects, an individual is said to look for common amounts in the payoffs, to drop (or "edit") those common amounts, and then compare the remaining distinct payoff terms in order to rank the prospects. This has implications for application of the original ("non-cumulative") version of prospect theory to our experiments. For example, the concavity calibration in Proposition 1 is based on the assumption that an agent prefers the certain amount $x + a$ to the lottery $\{x, p; x + b\}$ for all $x \in [m, M]$. But x is a common amount in the certain payoff, $x + a$ and both possible payoffs in the lottery $\{x, p; x + b\}$. If this common (or "reference point") amount x is edited, that is eliminated from all payoffs, then all comparisons are between the certain amount a and the single lottery $\{0, p; b\}$ and there remains no interval $[m, M]$ over which to calibrate. In this way, the editing component of the original version of prospect theory appears to immunize that theory to concavity calibration critique.

Does editing immunize the theory from being tested with data from our experiments? The answer to this question is "no," as can be seen by applying editing to the lotteries and certain payoffs used in the Calcutta experiment. If we perform editing by subtracting from all payoffs in each row the amounts that set the lower lottery payoff in all rows in the Option A column equal to 0 then the resulting comparison in *every* row is between the lottery $\{0, 0.5; 50\}$ and the certain payoff 20. Alternatively, if we perform editing by subtracting from all payoffs in each row the amounts that set the certain payoff in the Option B column equal to 0 then the resulting comparison in *every* row is between the lottery $\{-20, 0.5; 30\}$ and the certain payoff 0. Whichever way editing is applied it has the same implication: that an agent will view all rows in Table 4 as involving exactly the same choice and hence make the same decision. The data reveal that 77 percent of the subjects made choices that are inconsistent with this prediction. This percentage is much higher than the percentages of subjects who made choices that imply implausible large-stakes risk aversion with other decision theories. Therefore, although the original version of prospect theory with editing of common reference payoffs is immune to concavity calibration critique it is found to be a less plausible theory of risk-taking behavior than expected utility theory, rank-dependent expected utility theory, and cumulative prospect theory because it has the highest rate of inconsistency with data from the experiment.

6.5 Implications of the Data for Expected Value Theory

The choice faced by a subject in a row of Table 4 is between the Option A lottery $\{x-20, 0.5; x+30\}$ and the degenerate Option B lottery $\{x, 1; 0\}$. Since the expected value of an Option A lottery is $x+5$, a risk neutral agent will always choose Option A rather than Option B. The middle column of Table 1 reports that 77 out of 180 (or 43%) of the choices made by subjects were inconsistent with risk neutrality. Also 19 out of 30 (or 63%) of the subjects made at least one choice inconsistent with risk neutrality.

7. Magdeburg Experiment for Probability Transformation Theories

An experiment with one real payoff treatment was conducted at the MAX-Lab of the Otto von Guericke University of Magdeburg in February 2007. In this experiment, all payoffs were denominated in euros (€).

7.1 Experimental Design

Subjects were asked to make choices in each of the nine rows shown in Table 7. Row number i, for $i = 1, 2, ..., 9$, presented a choice between (a) a lottery that paid €40 with probability $i/10$ and €0 with probability $1 - i/10$ and (b) a lottery that paid €40 with probability $(i-1)/10$, €10 with probability $2/10$, and €0 with probability $1 - [(i-1+2)/10]$. In each row, a subject was asked to choose among Option A (the two outcome lottery), Option B (the three outcome lottery), and Option I (indifference). The subjects were presented with the instructions at the beginning of the session where the payment protocol of selecting one of the nine rows randomly for money payoff (by drawing a ball from a bingo cage in the presence of the subjects) was clearly explained to the subjects in the instructions as well as orally. The instructions also explained that if a subject chose Option I then the experimenter would flip a coin in front of the subject to choose between Options A and B for him (if that row was randomly selected for payoff). It was also explained that payoff from the chosen lottery would be determined by drawing a ball from a bingo cage in the presence of the subject. Appendix B.2 provides more information on the experiment protocol.

7.2 Implications of the Data for Dual Expected Utility Theory

There were in total 32 subjects in this experiment. For 8 subjects who switched at most once from Option A to Options B or I, and who choose Options B or I from row 4 (or earlier) to row 9, the dual theory predictions are reported in the top four rows of Table 8. Data for these eight subjects support the conclusion that the revealed risk aversion in the experiment implies large-stakes risk aversion for which $\{z, p; 0\} \succ \{t \times z, 0.5; 0\}$, that is, for which a lottery that pays z with probability p and pays 0 with probability $1 - p$ is preferred to a lottery that pays $t \times z$ or 0 with probability 0.5. Furthermore, the indicated preference holds for all positive values of z. For example, the risk aversion

revealed by subject 17 implies (from setting $z = \$4{,}000$) that he or she would prefer the lottery that pays $4,000 with probability 0.9 and pays $0 with probability 0.1 to the lottery that pays $324,000 or $0 with probability 0.5. Similarly, the data for subject 7 implies that he or she prefers the lottery that pays $4,000 for sure to the lottery that pays $976,000 or $0 with probability 0.5. The implied aversions to large-stakes risks reported in Table 8 for subjects 7, 17, and 22 are clearly implausible, while those for subjects 1, 8, 9, 13, and 21 are arguably implausible.

Another five subjects who chose Option B in row 1 and either Option B or I in row 9 reveal risk preferences that can be calibrated, as follows. We assume that if an individual switches from (risky) Option B choices to a (more risky) Option A choice, and then back to Option B, that the individual is less risk averse but not locally risk preferring at the switch row. In that case, data for the five subjects support the conclusion that an individual is predicted by dual expected utility theory to prefer a certain payoff in amount z to playing a 50/50 lottery with payoffs of 0 or the multiple of z reported in the bottom 5 rows of Table 8, where z is *any* positive amount. For example, the data for subject 11 support that conclusion that he or she would prefer receiving $\$4{,}000$ for sure to playing a fair bet with payoffs of $0 or $396,000. The implied aversions to large-stakes risks for subjects 4, 5, and 26 are arguably implausible, while those for subjects 11 and 30 are clearly implausible. Summarizing, the large-stakes risk aversion implied by the dual theory for 16% (five out of 32) of the subjects is clearly implausible whereas for another 25% (eight other subjects out of 32) of the subjects the implications are arguably implausible.

7.3 Implications of the Data for Expected Value Theory

The expected value of Option A in row i of Table 7 is €($40 - 4i$) while the row i Option B expected value is €($38 - 4i$). Hence a risk neutral agent will choose Option A over Option B in all rows. Column (4) of Table 1 reports that 175 out of 288 (or 61%) of the subjects' choices were inconsistent with risk neutrality. Also 31 out of 32 (or 97%) of the subjects made at least one choice inconsistent with risk neutral preferences.

8. Magdeburg Contingent Payoff (Casino) Experiment for Payoff Transformation Theories

An experiment was conducted in Magdeburg in the winter of 2004 with contingent money payoffs. This experiment was conducted at the MAX-Lab of the Otto von Guericke University of Magdeburg and the Magdeburg Casino. The subjects were adults who were older than typical students. They were recruited by announcements at the University of Magdeburg, in lectures in the adult program of the university and with letters to randomly-selected adult people in Magdeburg. All payoffs were denominated in euros (€).

8.1. Experimental Design

There were two sessions, one with 20 subjects and the other with 22. The experiment had three parts consisting of Step 1, Step 2, and a questionnaire. In Step 1 all subjects faced six decision tasks involving choices between a risky lottery $\{€x, 0.5; €y\}$, where $y = x + 210$, and the certain amount of money, $€z$, where $z = x + 100$ and where z took values from $(3K, 9K, 50K, 70K, 90K, 110K)$ and $K = 1,000$. Decision task 7 varied across subjects, depending on their choices in the first six decision tasks. In task 7, the choice was between the certain amount of money $9K$ and the risky lottery $\{€3K, 0.5; €G\}$, where G was chosen from $(50K, 70K, 90K, 110K)$ depending on the choices made by the subject in the first six decision tasks. The value of G was chosen so that if a subject had rejected the specified lotteries then the prediction from Proposition 1 was that he should reject the large-stakes lottery with the chosen G as well. Step 2 involved bets on an American roulette wheel at the Magdeburg Casino, the realization of which determined whether the euro payments determined in Step 1 were made in real euros.

Step 1 took place in the MAX-Lab in Magdeburg and lasted about 45 minutes. The participants made their decisions in well separated cubicles. First the instructions of Step 1 and the choices were given to the subjects. After the subjects completed Step 1 they got a questionnaire. After having completed the questionnaire they received the instructions for Step 2. In Step 1 the participants were told that their payoffs depended on a condition which would be described later in Step 2.

The decision tasks in Step 1 were choices between a binary lottery with probability $p = 0.5$ for each outcome, named Option A, and a sure payoff named Option B. The probability of $p = 0.5$ was implemented by flipping a coin. The subjects could choose one of the options or indifference. First,

subjects had to choose among Option A, Option B, and Option I (indifference) in each of six decision tasks corresponding to the six rows shown in Table 9. After they had finished this task the experimenter collected the response forms. Depending on the individual decisions of the subjects in decision tasks 1-6, an additional sheet of paper was handed out to the participants with one of the decision task 7 choices.

In Step 2 the payoff procedure was described. After the instructions for Step 2 were read by the participants, they were given the opportunity to change their decisions in Step 1. Nobody changed his decisions. Money payoff to a subject was conditional on an experimenter winning a gamble in the casino. Based on conditional rationality, all choices had the same chance to become real and the condition should not influence decisions.

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 Magdeburg 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 make it feasible to pay any of the amounts involved in the Step 1 decision tasks. If the casino bets placed for a subject paid off, one of that subject's Step 1 decisions would be paid in real euros, otherwise no choice would be paid. The decision that would be paid would be selected randomly by drawing a ball from an urn containing balls with numbers 1 to 7. The number on the ball would determine the decision task to be paid. If a subject had chosen indifference then a coin flip would determines whether the certain amount was paid or the lottery would be played. We informed the participants that any money resulting from casino bets that was not paid (because the subject's decision randomly selected for payoff involved amounts less than €110,250) would be used for subject payments in other experiments. Some more details of the protocol are explained in appendix B.3.

8.2 *Implications of the Data for Expected Utility Theory*

There were in total 42 subjects in this treatment. Eleven subjects never rejected a risky lottery. There were 20 subjects who revealed an interval of risk aversion with length at least 40K. Proposition 1

implies implausible large-stakes risk aversion for these 20 individuals, as reported in Table 10. Hence, we can conclude that 48% of the subjects satisfy the risk aversion assumption in intervals large enough to generate implausible risk aversion in the large for expected utility theory. The theory predicts that all 20 of these subjects should reject the constructed task 7 risky lotteries; 80 percent of these subjects did not reject the task 7 risky lotteries.

This experiment has no implications for cumulative prospect theory or rank dependent expected utility because, with the probability transformation function h(0.5) = 0.58, the lotteries $\{x, 0.5; x+210\}$ are not 100_D-favorable for either theory; that is inequality (3) is not satisfied and therefore Proposition (1) does not apply.

8.3 *Implications of the Data for Non-Cumulative Prospect Theory with Editing*

As explained in subsection 6.4 above, experiments of this type do have implications for the original version of prospect theory with editing. If we perform editing by subtracting from all payoffs on each row of Table 9 the amount that sets the lower lottery payoff in Option A equal to 0 then the resulting comparison in every row is between the lottery $\{0, 0.5; 210\}$ and the certain payoff 100. Alternatively, if we choose a different reference point and perform editing by subtracting from all payoffs in each row the amount that sets the certain payoff in Option B equal to 0 then the resulting comparison in every row of Table 9 is between $\{-100, 0.5; 110\}$ and 0. Both of these applications of editing (or other application that keeps the reference point fixed across rows) imply that an agent will make the same choice in every row of Table 9. The data reveal that 57% (24 out of 42) of the subjects made choices that are inconsistent with this prediction.

8.4. *Implications of the Data for Expected Value Theory*

The choice faced by a subject in a row of Table 9 is between an Option A lottery with expected value $x+105$ and a certain amount $x+100$. Therefore, a risk neutral agent will choose Option A in every row. Table 1 reports in column (5) that 125 out of the 252 (or 50%) of the choices made by subjects were inconsistent with risk neutrality. Also 31 out of 32 (or 97%) of the subjects made at least one choice inconsistent with risk neutrality.

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

Our finite St. Petersburg bet experiment, the Holt and Laury (2002) experiment, and a large number of other experiments in the literature support the conclusion that a large proportion of individuals are risk averse even in laboratory experiments. In addition, data for a large proportion of subjects in our calibration experiments are inconsistent with risk neutrality. Therefore, the expected value model does not adequately represent risk-taking behavior; a theory for risk averse agents is needed to model risky decision making.

Existing theories of risk aversion explain risk aversion with concave transformations of payoffs or pessimistic transformations of probabilities or transformations of both payoffs and probabilities. Such transformations, however, introduce issues of calibration of the implications of some patterns of subjects' choices for risk aversion in the large.

We report an experiment run in Calcutta, India with real rupee payoffs in amounts that were significant amounts of money to the subjects in the experiment. This is made clear by comparisons of the experiment payoffs to incomes received and prices paid by the subjects in their usual natural economic environment. Data from this experiment and the concavity calibration in Proposition 1 support the conclusion that a significant proportion of the subjects exhibit patterns of risk aversion that have implausible implications if one models their behavior with expected utility theory, cumulative prospect theory, or rank dependent expected utility theory. Although the original version of prospect theory with "editing" of reference payoffs can be immunized to problems from concavity calibration, the implications for predicted behavior of reference-point editing are inconsistent with data for most of the subjects.

Next, we report an experiment run in Magdeburg, Germany with real euro payoffs. Data from this experiment and the convexity calibration in Proposition 2 support the conclusion that a significant proportion of the subjects exhibit patterns of risk aversion that have implausible implications for risk aversion in the large according to dual expected utility theory.

Finally, we report an experiment run in Magdeburg with contingent payoffs of large amounts of euros. Data from these experiments and Proposition 1 imply implausible risk aversion in the large for about half of the subjects if one models their behavior with expected utility theory. Data for a

majority of subjects in this experiment are also inconsistent with testable predictions of the original version of prospect theory with editing of reference point payoffs.

Data from the experiments suggest that all existing theories of risky decisions may have implausibility problems. Data from our finite St. Petersburg game experiment, our calibration experiments, the Holt and Laury (2002) experiment, and many other experiments in the literature are inconsistent with expected value theory. Data from our calibration experiments are inconsistent with the original version of prospect theory with editing of reference point payoffs. Data for many subjects in our calibration experiments, together with Propositions 1 and 2, imply implausible risk preferences in the large for expected utility theory, cumulative prospect theory, rank dependent expected utility theory, and dual expected utility theory. Together, the theories examined in this paper represent all of the ways that risk aversion is conventionally modeled (i.e., by transforming payoffs and/or probabilities). Calibrations using data for many subjects imply implausible implications for all of the representative theories considered here, which suggests there may be no plausible theory of risky decisions.

Subsequent to completion of our experimental design, a new dual-self model (Fudenberg and Levine, 2006) was published that can rationalize various behavioral anomalies. The dual-self model can explain the paradox of risk aversion in the small and in the large from Rabin's (2000) thought experiment by use of a two-part utility function that has different risk preferences for small gains and losses ("pocket cash") than for large gains and losses (bank cash").

Fudenberg and Levine (2006, p. 1460) begin their explanation with a quotation from Rabin (2000): "Suppose we knew a risk-averse person turns down 50-50 lose \$100/gain \$105 bets for any lifetime wealth level less than \$350,000, but knew nothing about the degree of her risk aversion for wealth levels above \$350,000. Then we know that from an initial wealth level of \$340,000 the person will turn down a 50-50 bet of losing \$4,000 and gaining \$635,670." The dual-self model can rationalize these (thought experiment) outcomes when the \$100 loss and \$105 gain are "pocket cash" amounts and the \$4,000 loss and \$635,670 gain are "bank cash" amounts, as follows.[5] The \$100 loss and \$105 gain are evaluated with the pocket cash utility function (in Fudenberg and Levine's Theorem 2) using pocket cash as the reference "wealth parameter." This utility function implies rejection of the

50/50 lose $100 or gain $105 bet at an initial "wealth" level set equal to the pocket cash amount of $300 (a reference ATM daily withdrawal limit). The $4,000 loss and $635,670 gain in the large-stakes bet are too large to be pocket cash; they are bank cash amounts and hence evaluated with the bank cash utility function which uses actual wealth as the "wealth parameter." The bank cash utility function with initial wealth of $340,000 implies acceptance of the 50/50 lose $4,000 or win $635,670 bet. In this way, the dual-self model can rationalize the preferences to reject the small-stakes bet (at all actual wealth levels in a large interval) but accept the large-stakes bet (at a representative wealth level) in Rabin's (2000) thought experiment.

Can the dual-self model rationalize data from our experiments? First consider the choice options for the Calcutta experiment reported in Table 4. Recall that the experiment payoffs were denominated in rupees and that the subjects' monthly salaries were in the range 1,200 – 1,500 rupees. The payoffs in the first row of Table 4 are the same order of magnitude as two days' pay for the subjects, which might be considered pocket cash amounts. Payoffs in the other rows vary from a low of 980 rupees (65% - 82% of monthly salary) to a high of 6,030 rupees (4 - 5 months' salary). All of these payoffs are arguably bank cash amounts. Choice options in Table 9, for the Magdeburg "payoff transformation" experiment, involve payoffs that vary from 2,900 euros to 110,110 euros. These payoffs are all arguably bank cash amounts. Given that the choice options used in these experiments involve payoffs in bank cash amounts, the dual-self model has the same concavity-calibration implications as the familiar expected utility of terminal wealth model. In that case, the implausible risk aversion reported in Tables 6 and 10 that is implied by expected utility theory for choices by many of the subjects in our experiments cannot be avoided by applying dual-self arguments. In addition, the dual-self model is inconsistent with data from our experiments in a less fundamental way: the log utility function in the model (and other CRRA utility functions) imply acceptance of all of the bets in the Calcutta experiment and the Magdeburg "payoff transformation" experiment. This testable implication of the dual-self model is inconsistent with a large proportion of the data.

Endnotes

1. Subject instructions in English for all experiments reported in this paper are available on http://excen.gsu.edu.jccox.

2. The Calcutta, Casino and Dual experiments listed in Table 1 are explained below.

3. We use the term "utility functional" in a generic sense to refer to the functional that represents an agent's preferences over lotteries *in any decision theory*.

4. For the dual theory, in the case of binary lotteries, $h(p) = 1 - f(1 - p)$, where the function f is defined in statement (7) in Yaari (1987, p. 99).

5. The model can also rationalize the outcomes when the gain of $105 is considered bank cash (Fudenberg and Levine, 2006, p. 1460).

References

Bernoulli, Daniel, "Specimen Theoriae Novae de Mensura Sortis," *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, 5, 1738, 175-192. English translation: "Exposition of a New Theory on the Measurement of Risk," *Econometrica*, 22, 1954, 23-36.

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

Fudenberg, Drew and David K. Levine, "A Dual-Self Model of Impulse Control," *American Economic Review*, 96, 2006, 1449-1476.

Harrison, Glenn W., Eric Johnson, Melayne McInnes, and Elisabet Rutström, "Risk Aversion and Incentive Effects: Comment," *American Economic Review*, 95, 2005, 897 - 901.

Holt, Charles A. and Susan Laury, "Risk Aversion and Incentive Effects," *American Economic Review*, 92, 2002, 1644 – 1655.

Holt, Charles A. and Susan Laury, "Risk Aversion and Incentive Effects: New Data without Order Effects," *American Economic Review*, 95, 2005, 902-904.

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

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

von Neumann, John and Oscar Morgenstern, *Theory of Games and Economic Behavior*. Princeton University Press, Princeton NJ, 1947.

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

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*, 68, 2000, 1281-1292.

Roth, Alvin E., "Introduction to Experimental Economics," in John H. Kagel and Alvin E. Roth, *Handbook of Experimental Economics*. Princeton University Press, Princeton NJ, 1995.

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

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

Table 1. Observed Violations of Risk Neutrality

Nr of choices	St. Petersburg	Calcutta	Dual	Casino
Inconsistent with RN	127	77	175	125
Inconsistent with RN (%)	47	43	61	50
Total	270	180	288	252
Nr of subjects				
At least once not RN	26	19	31	32
At least once not RN (%)	87	63	97	76
total	30	30	32	42

Table 2. Concavity Calibrations for Payoff Transformations

Rejection Intervals (m = 900)	First EU Calibration	Second EU Calibration	First CPT & RDEU Calibration	Third EU Calibration	Second CPT & RDEU Calibration
M	**G**	**G**	**G**	**G**	**G**
5000	8,000	301,000	8,000	0.12×10^{17}	564,000
6000	10,000	0.15×10^{7}	10,000	0.4×10^{20}	0.29×10^{7}
8000	15,000	0.38×10^{8}	13,000	0.44×10^{27}	0.79×10^{8}
10000	24,000	0.98×10^{9}	18,000	0.49×10^{34}	0.21×10^{10}
30000	0.11×10^{9}	0.12×10^{24}	0.46×10^{7}	0.13×10^{105}	0.5×10^{24}
50000	0.1×10^{13}	0.14×10^{38}	0.34×10^{10}	0.37×10^{175}	0.1×10^{39}

Table 3. Convexity Calibrations for Probability Transformations

Unit division of [0,1] Interval	First Calibration c = 3	Second Calibration c = 4
Δ	**G**	**G**
0.1	33	244
0.05	1,025	59,050
0.01	$1,125\times10^{12}$	$7,179\times10^{2}$

Table 4. Choice Alternatives in Calcutta (Concavity-Calibration) Experiment

Option A	Option B	My Choice
80 or 130	100	A B I
980 or 1030	1000	A B I
1980 or 2030	2000	A B I
3980 or 4030	4000	A B I
4980 or 5030	5000	A B I
5980 or 6030	6000	A B I

Table 5. Calcutta Price Survey Data

Commonly used items for day-to-day living in Calcutta	Average Price range in Rupees
Food Items*	
Poultry	45-50
Fish	25-50
Red meat	150
Potatoes	7
Onions	10-12
Tomatoes	8-10
Carrots	8-10
Rice	11
Lentils	30
Public Transport	
Buses	
Local trains	3-4.5/ticket
	5-10/ticket
Eating out	15-35/person
Average restaurants	65-100/person
Expensive restaurants	500-1000/person
Five-star hotels/restaurants	

* Prices are in rupees per kilogram; 1 kilogram = 2.205 pounds

Table 6. Large-Stakes Risk Aversion Implied by Payoff Transformations for Calcutta Subjects

NOBS (30)	Observed Rejection Intervals (m, M)	G Values for EU $m+1K \succ_{EU} \{m, 0.5; G\}$	G Values for CPT & RDEU $m+1K \succ_{PT} \{m, 0.5; G\}$
2	(1K, 5K)	0.12×10^{17}	0.399×10^6
4	(2K, 6K)	0.12×10^{17}	0.4×10^6
2	(100, 4K)	0.54×10^{16}	0.338×10^6

Table 7. Choice Alternatives in Magdeburg Dual (Convexity-Calibration) Experiment

Row	Option A		Option B			Your Choice
	0 Euro	40 Euros	0 Euro	10 Euros	40 Euros	
1	1/10	9/10	0/10	2/10	8/10	A B I
2	2/10	8/10	1/10	2/10	7/10	A B I
3	3/10	7/10	2/10	2/10	6/10	A B I
4	4/10	6/10	3/10	2/10	5/10	A B I
5	5/10	5/10	4/10	2/10	4/10	A B I
6	6/10	4/10	5/10	2/10	3/10	A B I
7	7/10	3/10	6/10	2/10	2/10	A B I
8	8/10	2/10	7/10	2/10	1/10	A B I
9	9/10	1/10	8/10	2/10	0/10	A B I

Table 8. Large-Stakes Risk Aversion Implied by Probability Transformations for Magdeburg Subjects

Subject	Dual Theory Predictions
1, 8, 9, 13, 21	$\{z,0.7; 0\} \succ \{ 9z,0.5; 0\}$
22	$\{z,0.8; 0\} \succ \{27z,0.5; 0\}$
17	$\{z,0.9; 0\} \succ \{81z,0.5; 0\}$
7	$\{z, 1; 0\} \succ \{244z,0.5; 0\}$
4	$\{z, 1; 0\} \succ \{ 15z,0.5; 0\}$
5	$\{z, 1; 0\} \succ \{ 19z,0.5; 0\}$
11	$\{z, 1; 0\} \succ \{ 99z,0.5; 0\}$
26	$\{z, 1; 0\} \succ \{ 17z,0.5; 0\}$
30	$\{z, 1; 0\} \succ \{ 82z,0.5; 0\}$

Table 9. Choice Alternatives in Magdeburg Casino (Concavity-Calibration) Experiment

Option A	Option B	My Choice
2900 or 3110	3000	A B I
8900 or 9110	9000	A B I
49900 or 50110	50000	A B I
69900 or 70110	70000	A B I
89900 or 90110	90000	A B I
109900 or 110110	110000	A B I

Table 10. Large-Stakes Risk Aversion Implied by Payoff Transformations

for Magdeburg Subjects

NOBS	Observed Rejection Intervals (m, M)	G Values for EU $m+6K \succ_{EU} \{m, 0.5; G\}$
7	(3K, 110K)	0.21×10^{25}
1	(3K, 90K)	0.24×10^{21}
1	(3K, 50K)	0.3×10^{13}
8	(50K, 110K)	0.11×10^{16}
1	(50K, 90K)	0.13×10^{12}
2	(70K, 110K)	0.13×10^{12}

Appendix A. Proofs of Propositions and Corollaries

A.1 Proof of Proposition 1

Let a decision theory D with "utility functional" F_D in statement (1) be given, and let function φ be a concave non-decreasing function. Let $0 < a < b$ be given such that $\{0, p; b\}$ is an a_D-favorable risky lottery. Suppose that the agent prefers a positive amount of money $x + a$ for sure to a binary lottery, $\{x, p; x + b\}$ for all $x \in (m, M)$, $m > 0$. Let N be the largest integer smaller than *(M-m)/b*. This assumption and the definition of N imply

(a.1) $\qquad \varphi(x + a) \geq h(p)\varphi(x) + (1 - h(p))\varphi(x + b)$, for all $x \in (m, m + Nb)$.

First we show that (a.1) and concavity of φ imply that for all $z \in (m, m + Nb)$

(a.2) $\qquad \varphi'(z + jb) \leq \left(\dfrac{h(p)}{1 - h(p)} \times \dfrac{a}{b - a} \right)^j \varphi'(z)$, for all $j \in \Psi$,

where $\Psi = \{j \in \mathbb{N} \mid z + (j - 1)b \in (m, m + Nb)\}$.

$\qquad$ Next let K be the largest integer smaller than $(Z - m)/b$, and J be the smallest integer larger than $(\overline{G} - m)/b - K$ where $\overline{G}$ is the expression on the right hand side of inequality (*) in the statement of Proposition 1. We show that

(a.3) $\qquad \varphi(m + Kb) \geq h(p)\varphi(m) + (1 - h(p))\varphi(m + (K + J)b)$.

This completes the proof since all G that satisfy inequality (*) also satisfy $G < m + (K + J)b$, which together with (a.3) and the definition of K imply $\varphi(Z) > h(p)\varphi(m) + (1 - h(p))\varphi(G)$.

$\qquad$ To derive (a.2), first write $\varphi(x + a) = h(p)\varphi(x + a) + (1 - h(p))\varphi(x + a)$, next rewrite (a.1) with $x = z$, and finally group together terms with factors *h(p)* and *1-h(p)* on opposite sides of the inequality (a.1) to get

(a.4) $\qquad h(p)\big[\varphi(z + a) - \varphi(z)\big] \geq (1 - h(p))\big[\varphi(z + b) - \varphi(z + a)\big]$, $\forall z \in (m, m + Nb)$.

Inequalities $\big[\varphi(z + b) - \varphi(z + a)\big]/(b - a) \geq \varphi'(z + b)$ and $\big[\varphi(z + a) - \varphi(z)\big]/a \leq \varphi'(z)$, (both following from the concavity of φ) and inequality (a.4) imply

(a.5) $\qquad \varphi'(z + b) \leq \left(\dfrac{h(p)}{1 - h(p)} \dfrac{a}{b - a} \right)\varphi'(z)$, $\forall z \in (m, m + Nb)$.

Iteration of inequality (a.5) j times, for $j \in \Psi$, gives inequalities that together imply statement (a.2):

$$\varphi'(z + jb) \leq \left(\frac{h(p)}{1 - h(p)} \frac{a}{b - a} \right)\varphi'(z + (j - 1)b) \leq \ldots \leq \left(\frac{h(p)}{1 - h(p)} \frac{a}{b - a} \right)^j \varphi'(z).$$

$\qquad$ To show statement (a.3), let z denote $m + Kb$ and note that if $J + K > N$ then

$$\varphi(z+Jb)-\varphi(z)=\sum_{j=0}^{J-1}\left[\varphi(z+(j+1)b)-\varphi(z+jb)\right]$$

$$\leq b\left[(J-N+K)\varphi'(z+(N-K)b)+\sum_{j=0}^{N-K-1}\varphi'(z+jb)\right]$$

(a.6)

$$\leq b\varphi'(z)\left[\left(\frac{h(p)}{1-h(p)}\times\frac{a}{b-a}\right)^{N-K}(J-N+K)+\sum_{j=0}^{N-K-1}\left(\frac{h(p)}{1-h(p)}\times\frac{a}{b-a}\right)^{j}\right]$$

$$=b\varphi'(z)\left[q^{N-K}(J-N+K)+\frac{1-q^{N-K}}{1-q}\right]$$

(In (a.6) the first inequality follows from (weak) concavity of φ and $J+K>N$ whereas the second one follows from statement (a.2).) If however $J+K\leq N$ then one has

(a.6') $\qquad \varphi(z+Jb)-\varphi(z)\leq b\varphi'(z)\left(1-q^{N-K}\right)/(1-q)\leq b\varphi'(z)/(1-q)$

Similarly, one can show that

(a.7) $\qquad \varphi(z)-\varphi(z-bK)\geq b\varphi'(z)\sum_{k=0}^{K-1}\left(\frac{1-h(p)}{h(p)}\times\frac{b-a}{a}\right)^{k}=b\varphi'(z)\frac{(1/q)^{K}-1}{(1/q)-1}$

Hence, in case of $J+K>N$, (a.6) and (a.7) imply that a sufficient condition for (a.3) is

(a.8) $\qquad h(p)\frac{(1/q)^{K}-1}{(1/q)-1}\geq(1-h(p))\left[\frac{1-q^{N-K}}{1-q}+q^{N-K}(J-N+K)\right]$

or equivalently

(a.9) $\qquad J\leq N-K+\frac{1}{q^{N-K}}\left(\frac{h(p)}{(1-h(p))}\frac{(1/q)^{K}-1}{(1/q)-1}-\frac{1-q^{N-K}}{1-q}\right)=N-K+\frac{1}{1-q}+\frac{C}{b}q^{-N}$

The last inequality is true since

$$J\leq(\overline{G}-m)/b-K+1=\left(M+b\frac{2q-1}{1-q}+Cq^{-N}-m\right)/b-K+1$$

$$\leq\left(m+bN+b\frac{q}{1-q}+Cq^{-N}-m\right)/b-K+1$$

$$=N-K+\frac{1}{1-q}+\frac{C}{b}q^{-N}$$

Finally, if $J+K\leq N$, (a.6') and (a.7) imply that a sufficient condition for (a.3) is

(a.10) $\qquad \left(\left(\frac{1}{q}\right)^{K}-1\right)>\frac{1-h(p)}{h(p)}\frac{1}{q}$.

The last inequality is true since $q^{K-1} < h(p)$ which follows from the definition of K and by assumption $q^{<(Z-m)/b>-1} < h(p)$.

A.2. Proof of Proposition 2

First note that, according to dual expected utility theory, $R(i) \quad S(i)$, $i = 1,\ldots,2n-1$, implies

(a.10) $\qquad xf((1+i)\delta) + (c-1)xf((i-1)\delta) \geq cxf(i\delta), \ i = 1,\ldots,2n-1$

which is equivalent to

(a.11) $\qquad f((1+i)\delta) - f(i\delta) \geq (c-1)[f(i\delta) - f((i-1)\delta)], \ i = 1,\ldots,2n-1$

Writing inequality (a.11) for $i+k$ ($i+k = 1,\ldots,2n$) and reapplying it (k-1) other times one has

$$f((i+k)\delta) - f((i+k-1)\delta) \geq (c-1)[f((i+k-1)\delta) - f((i+k-2)\delta)] \geq \ldots$$
$$\geq (c-1)^k [f(i\delta) - f((i-1)\delta)]$$

which generalizes as

(a.12) $\qquad f(j\delta) - f((j-1)\delta) \geq (c-1)^{j-i}[f(i\delta) - f((i-1)\delta)], j = i,\ldots,2n$

Second, if we show that

(a.13) $\qquad f(0.5) \leq [f(0.5) - f(0.5-\delta)] \sum_{i=1}^{n} \left(\frac{1}{c-1}\right)^{i-1}$ and

(a.14) $\qquad 1 - f(0.5) \geq [f(0.5) - f(0.5-\delta)] \sum_{j=1}^{n} (c-1)^j$

Then we are done since inequalities (a.13) and (a.14) imply

$$\frac{1 - f(0.5)}{\sum_{j=1}^{n}(c-1)^j} \geq f(0.5) - f(0.5-\delta) \geq \frac{f(0.5)}{\sum_{i=1}^{n}(c-1)^{1-i}},$$

and therefore

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

To show inequality (a.13) note that $0.5 = n\delta$ and that

$$f(0.5) = \sum_{i=1}^{n}\left[f(i\delta) - f((i-1)\delta)\right] \le \left[f(n\delta) - f((n-1)\delta)\right]\sum_{i=1}^{n}\left(\frac{1}{c-1}\right)^{i-1}$$

$$= \left[f(0.5) - f(0.5-\delta)\right]\sum_{i=1}^{n}\left(\frac{1}{c-1}\right)^{i-1}$$

where the inequality follows from inequality (a.12). Similarly, inequality (a.14) follows from

$$1 - f(0.5) = \sum_{j=n+1}^{2n}\left[f(j\delta) - f((j-1)\delta)\right] \ge \left[f((n+1)\delta) - f(n\delta)\right]\sum_{j=1}^{n}(c-1)^{j-1}$$

$$\ge \left[f(0.5) - f(0.5-\delta)\right]\sum_{j=1}^{n}(c-1)^{j}$$

Appendix B. Additional Details of the Experiment Protocols

B.1   Calcutta Experiment for Payoff Transformation Theories

Each subject was asked to pick up a sheet of paper with either a number or a letter written on it. The subjects were presented with the instructions at the beginning of the session where the payment protocol of selecting one of the six tables randomly for money payoff (by rolling a six-sided die in the presence of the subject) was clearly explained to the subjects in the instructions as well as orally. The instructions also clarified that if they marked option I then the experimenter would flip a coin in front of the subject to choose between options A and B for him (if that decision was randomly selected for payoff). It was also clarified that if the subject chose the risky lottery in the selected decision task, then the lottery payment would be determined by flipping a coin in the presence of the subject.

Once the subjects finished reading the instructions they were given six sheets of paper, each containing one of the rows from Table 4, and were asked to mark their choices for each table and write the number/letter that they had picked up at the beginning of the experiment on top of each sheet. After all subjects were done with their decisions, task 2 was given to them, which consisted of filling out an income survey questionnaire. Again the subjects were asked to write the number/letter on the wealth questionnaires that they had picked up. A subject's responses were identified only by an identification code that that was the subject's private information in order to protect their privacy with respect to answers on the questionnaire. At the end of the two tasks, the experimenter went to an adjoining room and called each of the students privately for payment. For each subject, a die was rolled to decide the relevant payoff table. Further, if the subjects had marked the risky alternative in the selected table then a convention of paying the lower amount if the head came up and the higher amount if tails came up was announced to the student subject and incorporated. The student was asked to leave the questionnaire in a separate pile in order to protect privacy of responses.

B.2   Magdeburg Experiment for Probability Transformation Theories

Once the subjects finished reading the instructions they were asked to mark their choices on the response form and write the ID number/letter that they had picked up at the beginning of the experiment on top of each sheet. After all subjects were done with their decisions, task 2 was given to

them, which consisted of filling out a questionnaire. Again the subjects were asked to write the number/letter on the questionnaires that they had picked up. A subject's responses were identified only by an identification code that that was the subject's private information in order to protect their privacy with respect to answers on the questionnaire. At the end of the two tasks, the experimenter went to an adjoining room and called each of the students privately for payment. For each subject, a ball was drawn from a bingo cage containing balls numbered 1,2,…,9 to decide the relevant decision row and a ball was drawn from another bingo cage to determine the lottery payoff.

B.3 Magdeburg Contingent Payoff (Casino) Experiment for Payoff Transformation Theories

After step 1 was finished the questionnaire was handed out to the participants. Every participant could choose whether to answer the questionnaire or not. She was paid 10 euros if she answered it. Since all participants could only be identified by a code the answers to the questionnaire could not be attributed to a personally-identifiable individual, but only to the choices 1-7 she made. All participants filled out the questionnaire.

In Step 2 we selected three subjects randomly (in the presence of all of the subjects) to accompany the experimenter to the casino and verify that he bet the money as described above. After the visit to the casino, the experimenter and the three participants returned to the university and informed the remaining subjects about the results. If a participant would have won, we would have drawn the balls from an urn afterwards and correctly performed the coin flip. Step 2 was executed some hours later, on the same day as Step 1, after the casino opened. (As it turned out, none of the bets placed on a roulette wheel in the casino paid off.)