Asset Integration and Attitudes toward Risk

Theory and Evidence

Andersen, Steffen; Cox, James C.; Harrison, Glenn W.; Lau, Morten; Rutström, E. Elisabet; Sadiraj, Vjollca

Document Version
Final published version

Published in:
Review of Economics and Statistics

DOI:
10.1162/rest_a_00719

Publication date:
2018

Citation for published version (APA):
Andersen, S., Cox, J. C., Harrison, G. W., Lau, M., Rutström, E. E., & Sadiraj, V. (2018). Asset Integration and Attitudes toward Risk: Theory and Evidence. *Review of Economics and Statistics*, *100*(5), 816-830. https://doi.org/10.1162/rest_a_00719

ASSET INTEGRATION AND ATTITUDES TOWARD RISK: THEORY AND EVIDENCE

Steffen Andersen, James C. Cox, Glenn W. Harrison, Morten I. Lau,
E. Elisabet Rutström, and Vjollca Sadiraj*

Abstract—We provide evidence that choices over small-stakes bets are consistent with assumptions of some payoff calibration paradoxes. We then exploit the existence of detailed information on individual wealth of our experimental subjects in Denmark and directly estimate risk attitudes and the degree of asset integration. We discover that behavior is consistent with partial, rather than full, asset integration. The implied risk attitudes from estimating these specifications indicate risk premiums and certainty equivalents that are a priori plausible. This theory and evidence suggest one constructive solution to payoff calibration paradoxes.

I. Introduction

DEBATE surrounding theories of decisions under risk and uncertainty has renewed interest in the arguments of the utility function over event outcomes. The local measure of risk aversion proposed by Arrow (1971) and Pratt (1964) for expected utility theory (EUT) is based on terminal wealth being the argument. However, there is nothing in the axiomatic foundation of EUT that requires one to use terminal wealth as the argument. Vickrey (1945) used income instead of terminal wealth, von Neumann and Morgenstern (1944, 1953) were agnostic, and Luce and Raiffa (1957) discussed alternatives such as scalar amounts of terminal wealth or income or, alternatively, vectors of commodities. Arrow (1964), Debreu (1959), and Hirshleifer (1965) developed models in which the arguments of utility functions are vectors of contingent commodities.

The choice of arguments of the utility function can have significant consequences for the inferences one can plausibly draw from empirical estimates of risk attitudes. Many economics experiments present participants with gambles over relatively small stakes and find that such gambles are frequently turned down in favor of less risky gambles with smaller expected values: modest risk aversion is the general finding. If the argument of the utility function is terminal wealth, then some specific patterns of small-stakes risk aversion have implausible implications for preferences over gambles where the stakes are no longer small. One example from Rabin (2000) is that the expected utility of terminal wealth model implies that an agent who turns down a fifty-fifty bet of losing \$100 or gaining \$110, at all initial wealth levels between \$100 and \$300,000, will also turn down a fifty-fifty bet with a possible loss of \$2,000 even when the gain is as large as \$12 million if they have an initial wealth of \$290,000. However, if the argument of the utility function is not terminal wealth but rather the stakes offered in the gamble itself, or some other nonadditive aggregation of initial wealth and the stakes, implications of this assumed pattern of small-stakes risk aversion are no longer ridiculous (implausible) risk aversion (Cox & Sadiraj, 2006).

Given the importance of understanding the arguments of the utility function, the absence of empirical tests is remarkable. We initially provide evidence that choices over small-stakes bets in Denmark are consistent with suppositions in the payoff calibration paradoxes. We then present evidence from a unique data source that allows us to confront the question of whether integration of wealth with income in risk preferences is full, partial, or null when agents are making choices over gambles with more modest stakes. We combine field experimental data on lottery choices from a sample of the Danish population and individual-level information on personal wealth from a confidential database maintained by Statistics Denmark. Using these data we are able to identify a measure of personal wealth for the very same individuals who participated in standard experimental tasks. This allows us to explore theoretical specifications that measure the extent to which individuals integrate their wealth with the prizes on offer in the experimental lottery tasks.

We find no support for the terminal wealth model. We consider the evidence pooling over all subjects, assuming homogeneous preferences. Our subjects behave as if they integrate only a tiny fraction of their personal wealth with the lottery prizes they are asked to make choices over. In effect, this "weighted wealth" is indistinguishable statistically and economically from 0.[1]

In section II, we briefly review the theoretical literature on the arguments of utility over vectors of outcomes and implications for the measurement of risk attitudes. We note that calibration issues apply to a wide range of decision models. Moreover, extreme assumptions about the nature of asset integration can be seen as special cases of a more

Received for publication August 8, 2015. Revision accepted for publication October 2, 2017. Editor: Yuriy Gorodnichenko.

* Andersen: Copenhagen Business School; Cox: Georgia State University; Harrison: Georgia State University and University of Cape Town; Lau: Copenhagen Business School and Durham Unversity Business School; Rutström: Örebro University and Stockholm School of Economics; Sadiraj: Georgia State University.

We thank the U.S. National Science Foundation for research support under grants NSF/HSD 0527675, NSF/SES 0616746, and NSF/SES 0849590, and the Danish Social Science Research Council for research support under projects 24-02-0124 and 275-08-0289. We also thank Lasse Jessen, Jimmy Martínez-Correa, Bill Schworm, and three referees for helpful comments.

A supplemental appendix is available online at http://www.mitpressjournals.org/doi/suppl/10.1162/rest_a_00719.

---

[1] In an online appendix, we also consider the evidence for each subject individually, allowing each subject to have different risk preferences and different levels of asset integration. We find that 77% of our subjects behave as if they have a weighted baseline wealth of less than 10 kroner when evaluating risky lotteries, and 83% behave as if they have a weighted baseline wealth less than 1,000 kroner. None behave as if they fully asset integrate.

The Review of Economics and Statistics, December 2018, 100(5): 816–830

doi:10.1162/rest_a_00719

flexible specification that admits both wealth and experimental income as arguments of some nonlinear function. These results are not new, but they are not widely known. They are important because they serve up a menu of theoretically coherent alternatives to the extreme all-or-nothing assumptions about asset integration that are often subsumed in the literature.

In section III, we describe the data we have assembled from a combination of experimental tasks and links to Danish Registry databases maintained by Statistics Denmark (SD). The sense in which our measure of "personal wealth" deserves quotation marks is explained. It does not include everything that a theorist might want to see in there, such as the present subjective value of human capital, nor does it include every category of financial wealth. Nevertheless, it is arguably the most comprehensive wealth measure available to those who are interested in testing the theories of decision under risk.

In section IV, we present the structural model and econometric assumptions used to evaluate the extent of asset integration inferred from our data, and implications for risk attitudes. Section V presents estimates and implications. Section VI outlines some issues that arise in the general case in which experimental choices and nonexperimental choices are evaluated jointly.[2] Section VII draws conclusions.

We make two contributions. The first is to develop a general framework that clarifies that the core issue in the payoff calibration debate is the extent to which preferences are over income or final wealth, and it embeds the two extremes that have characterized the debate as special cases. The second contribution is to use this framework to estimate the extent to which risk attitudes to income variation in the lab are integrated with wealth using two sources of wealth variation. One source is a within-subject manipulation of variations in wealth and demonstrates that perfect asset integration would indeed be consistent with payoff calibration puzzles for our sample. The second source is cross-sectional variation of wealth in the Danish population, under the assumption of homogeneous preferences, and it implies very little asset integration. Experimental subjects' choices under the first source of wealth variation establish the need for examination of the second source of wealth variation.

II. Theory

A. Calibration Critiques

Some seemingly plausible patterns of small-stakes risk aversion can be shown, through concavity calibration arguments, to have implausible implications for large-stakes gambles under the terminal wealth specification, where initial wealth and income are integrated perfectly. Alternative empirical identifications of small-stakes patterns have implausible large-stakes implications for models defined on income in which there is no integration of wealth with income. A different type of (convexity) calibration analysis applies to models with nonlinear probability transformations.[3] From this literature, the theories that are now known to be subject to calibration critique include expected utility theory, dual theory, rank dependent utility (RDU) theory, cumulative prospect theory, and weighted utility and betweenness theories.

There are two types of calibration critiques that one needs to be cognizant of; we refer to these as payoff calibration critiques and probability calibration critiques. We consider the implications of the payoff calibration critiques. Within that category of critiques, the same risky (low-stakes) lottery choices can have quite different implications depending on the extent to which wealth is integrated with income in risk preferences. This is our principal focus, once we consider the empirical validity of the "seemingly plausible patterns of risk aversion" that underpin the calibration critique.

B. Small-Stakes Risk Aversion

The payoff calibration critique may be stated in terms of four suppositions:

- P = The agent is a risk-averse EUT maximizer.
- Q = The agent fully asset integrates.
- R = The agent (weakly) turns down small-stakes gambles in favor of a certain amount with a slightly lower expected value and does so over a large enough[4] range of wealth levels W.
- S = The agent turns down large-stakes gambles in favor of a certain amount with a significantly lower value, and looks silly.

The calibration puzzle is the claim that if P, Q, and R are true, then S follows. Since the behavior implied by supposition S is a priori implausible from a thought experiment, something must be inconsistent with these suppositions. Rabin (2000) draws the implication that P must then be false and that one should employ models of decision making under risk that relax supposition Q, such as cumulative prospect theory. As a purely logical matter, of course, this is just one way of many ways to resolve this calibration puzzle.

Evidence claimed to support the premise in statement R that decision makers in experiments exhibit small-stakes risk aversion for a large enough finite interval of wealth levels comes from designs in which subjects come to the experiment with potentially varying levels of wealth and each makes a single decision about a small-stakes lottery

---

[2] An online appendix reviews related literature in detail.

[3] See Hansson (1988), Rabin (2000), Neilson (2001), and Safra and Segal (2008) for concavity calibrations of terminal wealth models; Cox and Sadiraj (2006) and Rieger and Wang (2006) for payoff calibrations of income models; and Cox et al. (2013) and Sadiraj (2014) for probability calibrations of models with nonlinear probability transformations.

[4] The expression "large enough" is deliberately vague: it depends on the degree of risk aversion under supposition R and the lotteries in statement S that a priori exhibit silly behavior.

(Barberis, Huang, & Thaler, 2006). This is weak, indirect evidence, although it might be suggestive.[5] Interpretation of these data as providing a test of supposition R requires that we assume no variation of risk preferences between subjects and full asset integration (FAI) and accept guesses rather than data about wealth levels.[6] What is needed to evaluate supposition R is an experimental design that varies the wealth of a given decision maker who makes multiple decisions and can be presumed to behave consistently with one utility function during the lab session. Cox and Sadiraj (2008) propose a simple experimental design that does just this.

Cox and Sadiraj (2008) propose that one give subjects choices between a safe lottery of w for sure, and a risky lottery of a 50:50 chance of $w - x$ or $w + y$, where $w - x \geq 0$ and $y > x$. The key idea is to vary *w in the lab* and ask each subject to make lottery choice decisions at different levels of w. Consider values of w from the ordered set $\mathrm{S} = \{w, w, w, w, w, w, w\}$, where smaller font sizes of the letter w denote smaller values of lab wealth. These values of lab wealth may be plausibly much less than the $\boldsymbol{W}$ that the subject has in the field prior to the experiment. The experimenter does not need to know $\boldsymbol{W}$ for a given subject, but by varying "lab wealth" from S for that subject, the experimenter has considered small-stakes lottery choices over fifty-fifty probabilities of a low prize of $w - x$ and a high prize of $w + y$ against "lab wealth" w for sure, or "field + lab" wealth levels $\boldsymbol{W} + w$, with w from S, for that subject. This step of the design presumes that we vary lab wealth for a given subject, since then we can plausibly presume that field wealth $\boldsymbol{W}$ is constant for that subject during the experimental session. Integration of field wealth $\boldsymbol{W}$ with data from the experiment in analysis of calibration paradoxes depends on the existence of good data about field wealth and also assumes a version of supposition Q for which the agent perfectly asset integrates field wealth and lab wealth.

If the agent prefers the safe lottery over the risky lottery for all of the lab wealth values used in the experiment, then we have verification of supposition R, at least for the range of variation in wealth prescribed by the experimenter's budget. If we observe the agent choosing the safe lottery for small levels of lab wealth but the risky lottery for larger levels of lab wealth, then supposition R is rejected for that agent. Of course, we do not expect deterministic patterns of choice, so one ought to make some claim about the statistical significance of these choice patterns. This is one of the reasons for having multiple choices for each subject. An attractive feature of this experimental design is that we need not structurally model the EUT decision process for the agent; we can rely on simple statistical models such as (panel) probit, conditioned on lab wealth.

Building on this design, there have been "lab" tests of the premises of the calibration claims by Cox et al. (2013) and Harrison et al. (2017) that do not require integration of field wealth with lab wealth.

C. *Partial Asset Integration within EUT*

If supposition R cannot be rejected for the population under study, we must consider the implications of the payoff calibration critique in a constructive manner, and for that we turn to the idea of partial asset integration of wealth and income. We develop our analysis for a class of expected utility models that includes as special cases models with full asset integration (FAI), models with no asset integration (NAI), and models with partial asset integration (PAI). Models with full asset integration are possibly subject to the payoff calibration critique of Hansson (1988) and Rabin (2000). Models with no asset integration or partial asset integration are possibly subject to the payoff calibration critique of Cox and Sadiraj (2006) and Rieger and Wang (2006). Rather than engage in a priori arguments or thought experiments about paradoxes of risky choice, we develop a general theoretical model and let real data do some "real talking" in combination with that theoretical structure.

Cox and Sadiraj (2006) discuss the expected utility of initial wealth and income model with utility functional

$$\int u(w, y)\,\mathrm{d}G = \mathrm{E}_G(u(w, y)), \tag{1}$$

where G is an integrable probability distribution function and u is a utility function of initial wealth w and income y. We refer to this as the PAI-EUT model. Two standard models included in the PAI-EUT model are the expected utility of terminal wealth model with full asset integration (FAI-EUT), for which $u(w, y) = v(w{+}y)$, and the expected utility of income model with no asset integration (NAI-EUT), for which $u(w, y) = \xi(y)$.[7] These two standard models are polar cases in the class of models of PAI.

In our application, we treat w as deterministic and known, and of course y is stochastic by experimental design. This is consistent with the usual way in which asset integration is discussed in the literature. We discuss this issue further in section VI.

We begin with a quasi-concave utility function $u(w, y)$ defined over money payoff in the lab, y, and a measure of wealth, w. In a typical experiment, subjects' payoffs are in amounts of cash that may not be a perfect substitute for outside the laboratory wealth because of differences in liquidity and transaction costs. For example, \$100 in housing

---

[5] Schechter (2007) reports an experiment in which households each make one lottery choice and self-report their daily incomes. Interpretation of these data as a test of a calibration paradox requires maintaining the assumptions of no variation of risk preferences between households, linearity in intertemporal utility, and full asset integration.

[6] A common alternative assumption in the experimental literature is to assume no asset integration and interpret variation across wealth and observed choices across subjects as heterogeneity of risk preferences. It is apparent that both interpretations rest on previously untested, and extreme, assumptions about the degree of asset integration ("full" or "none," respectively).

[7] Any utility function of the form $u(w,y) = \xi(y) + h(w)$ would exhibit the same risk preferences over income y as does $\xi(y)$.

equity is not a perfect substitute for \$100 in cash received from participation in an experiment. Therefore, we consider the possibility that money payoffs in an experiment and wealth outside the laboratory may not be perfect substitutes.[8] There is then a need to distinguish curvature of indifference curves due to preferences over (w, y) from the preferences over risk.

D. Parametric Structure

A constant elasticity of substitution (CES) function can be used to aggregate wealth w and money payoff y when there is no risk. The terminal wealth model is found at one extreme of parameter values and the pure income model at the other. But the real interest is in between these extremes, and the point is to let the behavior of our subjects tell us the extent to which they (behave *as if* they) are integrating wealth with income from the experiment in making their choices.

Assume that all agents have the same ordinal preferences (when there is no risk), but can differ in their cardinal preferences (over risky outcomes).[9] We begin by studying homothetic preferences. Following Debreu (1976), there exists a least concave function, which is a cardinal utility that represents the same ordinal preferences. In case of homothetic preferences, the least concave function is a homogeneous function of degree 1, so we use the CES specification

$$v(w,y) = [\omega w^{\rho} + y^{\rho}]^{1/\rho} \tag{2}$$

where $w \geq 0$ is a measure of individual wealth, $y \geq 0$ is the prize in the money payoff in the experimental task, ω is a distributive share parameter to be estimated, $\sigma = 1/(1 - \rho)$ is the revealed elasticity of substitution between wealth and experimental money payoff, and is also to be estimated, and $-\infty < \rho \leq 1$ to ensure that $v(.)$ is quasiconcave. Risk-averse preferences over (w, y) are represented by concave transformations of this function and the EUT assumption that objective probabilities are not modified to generate decision weights. An often used specification of such transformation is the power function

$$U(v) = v^{1-r}/(1-r) \tag{3}$$

where $r \neq 1$ and v is defined by equation (2). In effect, equations (2) and (3) define a two-level, nested utility function, where equation (2) is an aggregator function defining a composite good, and equation (3) is the utility function defined directly over that composite.[10] Thus we can rewrite equation (3) more compactly as

$$U(w,y) = \left[(\omega w^{\rho} + y^{\rho})^{(1-\mathrm{r})/\rho}\right]/(1-\mathrm{r}), \tag{4}$$

where $\omega\ w^{\rho} + y^{\rho} > 0$.[11] This generalized CES function blends together full, partial, and null asset integration on (w, y) space with risk preferences on composite good, $v(w, y)$, space.

With these parametric assumptions, the familiar one-dimensional Arrow-Pratt measure of relative risk aversion with respect to y, evaluated at w, is then

$$[r\,y^{\rho} - w^{\rho}(\varrho - 1)\omega]/[y^{\rho} + w^{\rho}\omega]. \tag{5}$$

We discuss the need for measures of multivariate risk aversion in section VI if one is to generalize our approach to allow both arguments of the utility function to be random.

Perfect asset integration with the utility of terminal wealth EUT model is the special case in which $\omega > 0$ and $\sigma = \infty$. The usual case in the literature assumes further that $\omega = 1$, so that income and wealth are added together

---

[8] It is the case that if w and y are allowed to be imperfect substitutes, then we have to assume the possibility of imperfect markets in w and y, or else some elementary no-arbitrage conditions would be violated. We do not view this as particularly problematic for three reasons. First, if behavior is better characterized by assuming that w and y *are* indeed imperfect substitutes, then we have to assume imperfect markets. But then that assumption is one that is in effect supported by the data, even if it runs counter to some stylized model of behavior. That is, imagine that w and y are imperfect substitutes in preferences but perfect substitutes at some relative price in the market. Then we would never observe behavior suggesting that they are perfect substitutes; hence, we would never observe full asset integration behavior. The second reason that we do not view the assumption of imperfect markets as problematic is that there are transactions costs in converting one asset to another, at least for the assets we consider. These transactions costs might be larger or smaller for different individuals or for different asset classes when one considers generalizations (as we do in section VI), but those have to be evaluated on a case-by-case basis. The third reason is related to the second: we could imagine an even more general model in which the degree of asset integration emerges endogenously as a function of circumstances; these could be the transactions costs faced in substituting assets in the market, but it could also be the cognitive burden of thinking of the assets as perfect substitutes in preferences. That is, for some unstated reason, the agent might prefer to keep w and y in distinct mental accounts, but still think of them as substitutable to some degree.

[9] In a univariate model with either income or wealth as the only argument, cardinality is modeled entirely through the concavity of the utility function over the single argument. Here, however, cardinality depends also on the convexity of the contour functions over the two imperfectly substitutable utility arguments.

[10] This power function is unbounded, so it is useful to be clear on the implications for concavity calibration puzzles under FAI and EUT on a bounded or unbounded domain. If the utility function is bounded on $(0, \infty)$, then that is a sufficient condition for implausible risk aversion in large stakes (e.g., Cox & Sadiraj, 2008, proposition 2); global small-stakes risk aversion is not needed for this result. It is not a necessary condition. Small-stakes risk aversion over all $(0, \infty)$ is a sufficient condition for the utility function to be bounded (e.g., Rabin, 2000, or Cox & Sadiraj, 2006); it is not, however, a necessary condition. Being bounded on $(0, \infty)$ is a necessary condition for small-stakes risk aversion over the open interval $(0, \infty)$, but it is not sufficient. An increasing power function is unbounded and hence violates the necessary condition on boundedness; therefore, it cannot represent risk attitudes that exhibit small-stakes risk aversion over all $(0, \infty)$. The sufficiency part can be illustrated by considering a constant absolute risk aversion (CARA) function with parameter 0.0003. It is bounded; however, the small-stakes risk aversion pattern in Cox and Sadiraj (2006) is not satisfied, since \$100 for sure is rejected in favor of an equal chance of \$210 or \$0. Small-stakes risk aversion defined on a finite interval implies nothing at all about the boundedness of the utility function. Finally, small-stakes risk aversion over a large enough finite interval is a sufficient condition for implausible risk aversion for large stakes, whether or not the utility function is bounded or unbounded.

[11] For negative prizes in income, write it as $\omega\ w^{\varrho} + \mathrm{sign}(y)\ \mathrm{abs}(y)\ \rho > 0$.

on a one-to-one basis. Zero asset integration with the utility of income EUT model, where income is interpreted tightly to mean the income from this specific experimental choice,[12] is the special case in which $\omega = 0$.[13] Note that we say nothing in this case about σ, because any value of σ would generate the same observed choices if $\omega = 0$. Our null hypothesis is that subjects perfectly asset integrate with their actual wealth.

III. Data

Our data consist of observations of choice behavior in experimental tasks and wealth data for 442 individuals. The sample is representative of the adult Danish population residing in Greater Copenhagen as of January 2015. Our sample consists of 52% men, aged 47 on average, 43% of whom were married, with an average household size of 1.4, and with average income of 434,085 kroner per year. Compared to the 1,455,772 comparable Danes in the Registry, our subjects are not statistically significantly different except for household size and income: the population averages were 1.54 and 338,859 kroner, respectively.

All experiments were run in February and March 2015. The experimental data are of the standard type and employ procedures described in Andersen et al. (2014).

The wealth data are novel and involve matching the experimental subjects with data collected by SD. The matching process, and all statistical analyses with those data, occur remotely at the statistical agency, to ensure privacy.

A. *Experimental Data*

Each of our 442 subjects was asked to make choices for each of sixty pairs of lotteries in the gain domain, designed to provide evidence of risk aversion as well as the tendency to make decisions consistently with EUT or RDU models.[14] The online appendix lists these lottery parameters and the logic behind them. In general, each lottery has three prizes, although there are some lotteries with four prizes, two prizes or just one prize. The battery is based on ingenious designs from Wakker, Erev, and Weber (1994), Loomes and Sugden (1998) and Wilcox (2015), as well as the direct test of supposition R proposed by Cox and Sadiraj (2008) reviewed earlier. The analysis of risk attitudes given these choices follows Harrison and Rutström (2008).

There were four batteries used across the 442 subjects. Each battery included the 24 lottery pairs from Wakker et al. (1994). One battery also included 36 lottery pairs from Wilcox (2015), and this full set of 60 lottery pairs was administered to 222 subjects. The remaining three batteries included the lotteries inspired by Loomes and Sugden (1998) and Cox and Sadiraj (2008), for another set of 60 lottery pairs administered to 220 subjects; the three versions of this battery differed by varying the scale of payoffs.

We carefully selected these lotteries to ensure considerable variation in prizes and probabilities to facilitate identification of the full structural model. Over all batteries, there are 90 distinct prizes and 16 distinct probabilities. At the individual subject level, the number of distinct prizes is either 37 or 26, and the number of distinct probabilities is either 16 or 12.

Apart from the tests of supposition R, these choices themselves are not the direct basis for our evaluation of the payoff calibration paradoxes. Combined with the wealth data for each subject, these choices allow us to estimate the risk preferences implied by EUT and RDU models, and those estimates are then used to evaluate the paradoxes with counterfactual lottery choices. The many variations in wealth, lottery payoffs and lottery probabilities implied by our design allow us to identify all the required theory parameters.

B. *Wealth Data*

Wealth data are based on register data from SD. Our data contain economic, financial, and personal information on each individual from relevant official registers. The data set was constructed based on two sources made available from SD and matched with our experimental data; these sources are the Danish Civil Registration Office and the Danish Tax Authorities. All permanent residents in Denmark and all Danish citizens have a unique social security number given at birth or the date of formal residence, known as the CPR number, and this number allows us to match data across data sources. The CPR number follows every individual throughout the person's lifetime, and all information on an individual is registered on this number. We had access to the CPR number of every subject in our experiments.

Individual and family data are taken from the records in the Danish Civil Registration. These data contain the entire Danish population and provide unique identification across individuals and households over time. Each record includes the personal identification number (CPR), name, gender, date of birth, as well as the CPR numbers of nuclear family members (parents, siblings, and children) and marital history (number of marriages, divorces, and widowhoods).

---

[12] This interpretation is tight in the sense that one might also consider income from the set of experimental tasks that this binary choice is embedded in or the income from the whole experimental session. For example, is income the lottery prize in one binary choice pair, the income from the sixty choices, or the income from the whole session since there were additional paid choices in addition to these lottery choice questions? One could undertake an exactly parallel discussion of partial asset integration within the experimental session, evaluating what might be called local asset integration issues. Our focus here is on global asset integration issues between the usual interpretations of experimental data and the implications of the calibration critiques.

[13] To visualize these intuitively as perfectly complementary Leontief preferences, one might further assume $\sigma = 0$. This assumption, although often implicit, is not necessary for NAI.

[14] The subjects were also presented with other decision tasks in the experiment, which are not analyzed here. For each type of decision task, the subjects had a 10% chance of getting paid. If he or she was paid in the part of the experiment analyzed, one of the sixty decision tasks was randomly selected, and the chosen lottery was played out for payment. Average earnings for those who got paid from these sixty decision tasks was 1,923 kroner. Average earnings including recruitment fees across all 442 subjects was 954 kroner.

In addition to providing extra control variables, such as age, gender, and marital status, these data enable us to identify the subjects who participated in the artifactual field experiment described above.

Income and wealth information are retrieved from the official tax records at the Danish Tax Authorities (SKAT). This data set contains personal income and wealth information by CPR numbers on the entire Danish population. SKAT receives this information directly from the relevant sources: financial institutions supply information to SKAT on their customers' deposits, financial market assets, interest paid or received, and security investments and dividends. Employers similarly supply statements of wages paid to their employees.

The wealth variable in our analysis is constructed from data reported by SD that represent net individual wealth.[15] Total assets are the market value of domestic real estate, shares and mutual funds, bonds, assets deposited in domestic and foreign financial institutions, pensions, and the value of automobiles. Total liabilities are the value of debt in domestic and foreign financial institutions and mortgages. All values of shares, bonds, and pensions are reported by financial institutions as of December 31, 2014; values of real estate are estimated by SD as the market value on December 31, 2014; and the value of automobiles is calculated by SD with a one-year lag.[16] All values are in 2015 Danish kroner, and values are reported for the full sample of 442 subjects. (Conversions to USD use the exchange DKK 1 DKK = USD 6.643 applicable during most of the experiment.)

Our wealth measure does not include cash, value of yachts, paintings, equity in privately held companies, or the market value of shareholder equity in privately held companies and unlisted mutual funds. Our wealth measure does include shareholder equity in publicly traded companies and listed mutual funds. The wealth measure does not include nontraded assets such as human capital, which means that borrowing for assets such as education is seen as debt without any corresponding assets. This is arguably one of the most comprehensive measures of private financial wealth for an entire population that one can get, although we realize that some important nonfinancial components are left out.

Table 1 provides a tabulation of wealth and its components for our sample. The positive skew of the distribution of wealth is no surprise. For 4.7% of our subjects, or 21 out of 442, there is negative net wealth, reflecting the fact that some assets are not fully accounted for. For all calculations, we assume that wealth cannot be negative and truncate it to 0. Individuals with 0 field wealth have nothing to integrate with lab income, though in a formal sense, of course, they do integrate, but the effect is as if they do not since they have 0 wealth.

TABLE 1.—INDIVIDUAL WEALTH IN DENMARK

Variable	Mean	Median	SD
Total assets	3,844,104	2,985,522	4,521,335
Real estate	1,427,395	1,000,828	2,734,828
Shares and mutual funds	185,023	2,859	562,243
Bonds	4,006	0	28,118
Assets in financial institutions	186,747	65,762	311,192
Pensions	1,969,176	1,162,490	2,504,648
Automobiles	71,758	27,400	105,166
Total liabilities	769,426	352,192	2,212,928
Debt in financial institutions	190,558	26,133	439,769
Mortgages	578,869	0	2,023,922
Net wealth	3,074,678	2,165,847	3,470,853
Net wealth truncated at 0	3,097,435	2,165,847	3,439,401

All currency values in Danish kroner (DKK 1 = US $6.643 in September 2015). All valuations as of December 31, 2014, except for automobiles, which has a one-year lag. Total assets are the market value of domestic real estate, shares and mutual funds, bonds, assets deposited in domestic and foreign financial institutions, pensions, and the value of automobiles. Total liabilities are the value of debt in domestic and foreign financial institutions and mortgages. All values of shares, bonds, and pensions are reported by financial institutions as of December 31. Values of real estate are estimated by Statistics Denmark as the market value on December 31. The value of automobiles is calculated with a one-year lag. All foreign assets and debt are self-reported and equal to 0 for every subject in the sample. All values are in 2015 Danish kroner, and values are reported for the full sample of 442 subjects.

Access to these unique data is an important issue in terms of both the ability of others to replicate our findings and for their ability to extend our analysis. Researchers at authorized Danish institutions can gain access to deidentified microdata provided by SD through remote access connections. SD manages most of Danish microdata. The fundamental authorization principle of SD is that data will not be disclosed where there is an imminent risk that an individual person or individual enterprise can be identified. This applies not only to identified data, such as CPR numbers, but also to deidentified data, since such data are usually so detailed that identification can be made. The online appendix documents procedures to access these data.

IV. Econometric Model

A. *Expected Utility Theory*

Although the concerns about implausible risk attitudes under terminal wealth specifications apply to all decision theories that are additive over states, we initially focus on EUT because it is parsimonious. Under EUT, the probabilities for each outcome y_j, $p(y_j)$, are those that are induced by the experimenter, so expected utility is simply the probability-weighted utility of each outcome in each lottery $i \in \{A, B\}$, where A and B denote left and right lottery, respectively. Using $U(w, y)$ from equation (4), we then have

$$EU_i = \Sigma_{j=1,J}\left[\left(p\left(y_j\right)\right) \times U\left(w, y_j\right)\right] = \Sigma_{j=1,J}\left[p_j \times U\left(w, y_j\right)\right] \tag{6}$$

for a lottery with J prizes. To capture behavioral errors, we employ a Fechner specification with contextual utility, so that we assume the latent index

---

[15] An alternative is to use household wealth rather than individual wealth, exploiting further the ability of our data to identify other members of the household of the subject in our experiments. On the other hand, one then opens up subtle issues about whose risk attitudes were on display in the experiments (i.e., those of the individual or those of the household) and how households pool income from individuals.

[16] All foreign assets and debt are self-reported to SKAT and are 0 for every subject in our sample.

$$\nabla EU = [(EU_B - EU_A)/\tau]/\mu, \tag{7}$$

where τ is a normalizing term described in a moment, μ is the Fechner behavioral error parameter to be estimated, and EU_B and EU_A are the expected utilities of the right and left lottery as presented to subjects. The normalizing term τ is defined as the difference between the maximum utility over all of the prizes in that lottery pair minus the minimum utility over all of the prizes in that lottery pair. Thus, it varies from choice context to choice context, depends on the parameters of the utility function, and normalizes the difference in EU to lie between 0 and 1. This results in a more theoretically coherent concept of risk aversion when one allows for a behavioral error such as with μ (Wilcox, 2011).

The latent index, equation (7), based on latent preferences, is then linked to the observed experimental choices using a standard cumulative normal distribution function $\Phi(\bigtriangledown EU)$. This probit function takes any argument between $\pm\infty$ and transforms it into a number between 0 and 1 using this familiar function. Thus, we have the probit link function:

$$\text{prob}(\text{choose lottery B}) = \Phi(\nabla EU). \tag{8}$$

The index defined by equation (7) is linked to the observed choices by assuming that the probability that the B lottery is chosen depends on $\bigtriangledown EU$ in the manner specified by equation (8).

Thus, the likelihood of the observed responses, conditional on the EUT and utility specifications being true, depends on the estimates of the utility function given the above statistical specification and the observed choices. The log likelihood for the utility function, equation (4), is

$$\begin{aligned}\ln L(r, \omega, \rho, \mu; c, w) = &\sum_i [\,(\ln \Phi(\nabla EU) \times \mathbf{I}(c_i = 1)) \\ &+ (\ln \Phi(1 - \nabla EU) \times \mathbf{I}(c_i = -1))],\end{aligned} \tag{9}$$

where $\mathbf{I}(\cdot)$ is the indicator function, $c_i = 1(-1)$ denotes the choice of the option B (A) lottery in risk aversion task i, and $\bigtriangledown EU$ is defined using the parameters r, ω, ρ, and μ.[17] All estimates employ clustering at the level of the individual, since errors for a given individual may be correlated.

B. Rank-Dependent Utility Theory

One popular alternative to EUT is to allow the decision maker to transform the objective probabilities presented in lotteries and use these weighted probabilities as decision weights when evaluating lotteries. To calculate decision weights under RDU, one replaces expected utility defined by equation (6) with *RDU*:

$$RDU_i = \sum_{j=1,J} [(d(y_j)) \times U(w, y_j)] = \sum_{j=1,J} [d_j \times U(w, y_j)], \tag{10}$$

where

$$d_j = \pi(p_j + \ldots + p_J) - \pi(p_{j+1} + \ldots + p_J) \tag{11a}$$

for $j = 1, \ldots, J - 1$, and

$$d_j = \pi(p_j) \tag{11b}$$

for $j = J$, with the subscript j ranking outcomes from worst to best, $\pi(\cdot)$ is some probability weighting function, d_j is the decision weight on the jth-ranked outcome, and RDU refers to the rank-dependent utility model. Of course, one then has to specify the functional form for $\pi(p)$ and estimate additional parameters, but the logic extends naturally.

We use the general functional form proposed by Prelec (1998) for probability, since it exhibits considerable flexibility. This function is

$$\pi(p) = \exp\{-\eta(-\ln p)^{\varphi}\}, \tag{12}$$

and is defined for $0 < p \leq 1$, $\eta > 0$, and $\varphi > 0$. Note that we do not require $0 < \varphi < 1$. When $\varphi = 1$, this function collapses to the familiar power function $\pi(p) = p^{\eta}$, and EUT assumes the identity function $\pi(p) = p$, which is the case when $\eta = \varphi = 1$. With equation (12) included, the log likelihood then becomes

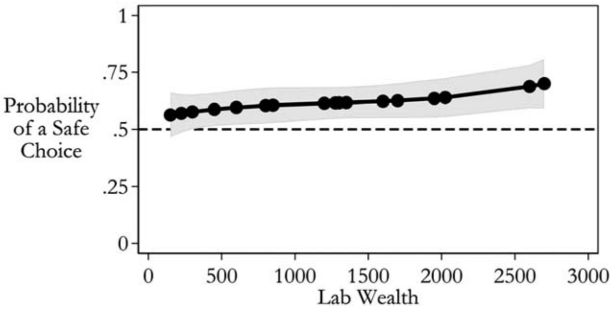
$$\begin{aligned}\ln L(r, \omega, \rho, \eta, \varphi, \mu;\ c,\ w) = &\sum_i [(\ln \Phi(\nabla RDU) \\ &\times \mathbf{I}(c_i = 1)) + (\ln \Phi(1 - \nabla RDU) \\ &\times \mathbf{I}(c_i = -1))],\end{aligned} \tag{13}$$

and we estimate the model with two extra parameters for the probability weighting function.

Estimating the RDU model from experiments that employ the random lottery incentive method (RLIM) requires that one assumes that individuals isolate each pairwise lottery choice within the series from each other. This implies the compound independence axiom, even though the RDU model allows independence to be violated when subjects evaluate each simple lottery. The vast majority of incentivized lottery choice experiments use RLIM and rely on this axiom. Thus, the RDU model applied to RLIM data inconsistently relaxes that axiom when it comes to evaluating individual lotteries, but assumes that it is valid when

[17] One of the core hypotheses to be tested is that $\omega = 0$, and one can run into issues with such hypothesis tests where the parameter in question is close to the boundary of an admissible region. In fact, we are estimating a likelihood function that is already highly nonlinear in the parameters (e.g., the curvature of the utility function). Hence, we can use a standard numerical method to constrain parameters such as ω to lie in the unit interval by estimating a different parameter $\tilde{\omega}$, which is then, in the function evaluator, converted to $\omega = 1/(1 + \exp(\tilde{\omega}))$. In this manner, the algorithm evaluating the likelihood can vary $\tilde{\omega}$ between $\pm\infty$ and still keep ω constrained to the unit interval. All hypothesis tests defined over ω are numerically undertaken on the estimated parameter $\tilde{\omega}$, which by definition never gets close to a boundary (the hypothesis tests are therefore nonlinear in nature and use the delta method to correctly infer test statistics and p-values).

FIGURE 1.—PREDICTED PROBABILITY OF A SAFE CHOICE BY ADULT DANES AT VARYING LEVELS OF LAB WEALTH



N = 220 subjects each making six choices for varying lab wealth. Predictions from a random effects panel probit model.

applying the RLIM payment protocol (Harrison & Swarthout, 2014; Cox, Sadiraj, & Schmidt, 2015).

V. Results and Implications

A. *Tests of the Small-Stakes Risk Aversion Premis*

Using the test proposed by Cox and Sadiraj (2008) for a subsample of 220 adult Danes from our complete sample of 442, we find evidence of the relevant type of small-stakes risk aversion for the range of lab wealth we considered. The experimental design involved them each making six binary choices in the wider battery of binary choices we consider below. Subjects were randomized to six lottery choice pairs from a set of 18 possible pairs, spanning 17 different levels of lab wealth. Hence, the lab wealth varied for each subject over their 6 choices, and we have pooled data spanning the 17 lab wealth levels. The gains and losses in absolute value were paired for each subject over different lab wealth levels—for example, +180 and −160 for lab wealth levels of 300 (≈\$45) and 2,700 (≈\$406). Although we refer to "lab wealth," all that the subject saw was a lottery that had one outcome with a probability of 1 and another lottery with the usual risky outcomes.[18] Hence, we did not use language or framing that would lead subjects to be more or less inclined to integrate it into their extra-lab "field wealth." Nor were the outcomes in the risky lottery presented as deviations from the certain outcome of the nonrisky lottery, which might also encourage framing. For example, subjects were asked to choose between 2,700 kroner for sure and the risky lottery with outcomes 2,540 kroner and 2,880 kroner. With perfect asset integration, 2,540 kroner adds to the subject's wealth no matter what the subject's choice is, so we refer to 2,540 kroner as "lab wealth."

Figure 1 shows the findings with a random-effects panel probit model, since there is no need here for structural estimation of risk preferences. We find no significant evidence of a decline in risk aversion for lab wealth levels over the range considered here. The solid line shows the average prediction, and the shaded area shows the 95% confidence interval around that prediction. Subjects exhibit risk aversion for all levels of lab wealth considered here, so we conclude that the evidence for these adult Danes and these levels of lab wealth does not lead us to reject supposition R, that "the agent turns down small-stakes gambles in favor of a certain amount with a slightly lower expected value, and does so over a large enough range of wealth levels W."

Since supposition R, one of the premises of the calibration critique, is not rejected, there is a need to examine the partial asset integration specification proposed earlier.

B. *Basic Results on Asset Integration for Representative Agents*

We now employ the full sample of 442 Danes and all of the 60 binary choices each of them made. Panel A of table 2 shows maximum likelihood estimates of the utility function, equation (4). We assume here that every adult Dane in our sample has the same ordinal preferences over w and y (when there is no risk), as well as the same coefficient r; the online appendix considers estimates for each individual. The coefficient r is estimated precisely, as is the parameter ω reflecting the weight attached to wealth. We find that the weight attached to wealth is virtually 0 and statistically not different from 0. This is a fundamental result, since it means that the PAI specification collapses to the NAI specification in this pooled estimation, and we reject the FAI hypothesis. It also means that it is virtually impossible, for sensible economic reasons, to identify the substitutability between w and y. We find an estimate of ρ of 0.89, implying an estimate of σ of 9.1, but since there is virtually no weighted wealth to substitute with, these values have little economic meaning.

Average net wealth in the estimation sample is 3,074,678 kroner (≈\$462,845), so these estimates imply that individuals behave as if they evaluate experimental income relative to a weighted baseline wealth of $\omega \times w = 3{,}074{,}678 \times 0.00000625 = 19$ kroner (≈\$2.86). This is effectively 0 in economic terms. For example, it would currently get only half of an Egg McMuffin Value Meal in a Danish McDonald's. Another way to evaluate this weighted baseline wealth estimate of 19 kroner is by comparison with the lottery prizes, which ranged between 0 kroner and 6,750 kroner (≈\$1,016). Needless to say, we can easily reject the hypothesis of FAI since $\omega \approx 0$, and the p-value on the test of the hypothesis that $\omega = 0$ is 0.77.

Another way to see these results, perhaps more intuitively, is to see if measures of net wealth correlate with risk attitudes in a reduced-form manner. We do this by estimating the EUT-NAI model and asking if the coefficient r is significantly affected by net wealth. In this case, we model r as a linear function of some covariates. Our structural

---

[18] An alternative way to add a lab wealth component might be to randomly add it to the show-up fee for participating in the experiment. The problem with this approach is that it would raise a potential confound due to sample selection issues.

Table 2.—Estimates and Implied Certainty-Equivalents Using EUT-PAI Model

A. Estimates

Parameter	Point Estimate	Standard Error	p-value	95% Confidence Interval	
r	0.64	0.04	<0.001	0.57	0.71
ρ	0.89	0.15	<0.001	0.6	1.19
ω	0.000006	0.00002	0.77	−0.00004	0.00005
μ	0.08	0.005	<0.001	0.07	0.09

B. Certainty-Equivalent Calculations with Average Wealth

High Prize (DKK)	Probability of High Prize	Low Prize (DKK)	Expected Value (DKK)	Certainty Equivalent (DKK)	Ratio
200	0.5	100	150	145	0.965
500	0.5	100	300	252	0.84
1,000	0.5	100	550	402	0.73
2,000	0.5	100	1,050	663	0.631
5,000	0.5	100	2,550	1,350	0.529
5,000	0.01	100	149	109	0.732
5,000	0.1	100	590	214	0.362
5,000	0.3	100	1,570	626	0.399
5,000	0.7	100	3,530	2,459	0.697
5,000	0.9	100	4,510	4,025	0.892

Sample of 442 individuals making 26,520 choices of strict preference. Log likelihood = −17,025 (−17,028 for NAI and −17,436 for FAI). Null hypothesis for p-value results is that the coefficient estimate is 0.

results suggest that they should not, since net wealth is zeroed out by a very low estimate of ω, at least when we assume homogeneous risk preferences. If we include net wealth, the effect on r is -0.004 with a p-value of 0.45; if we include a dummy for the top quartile of net wealth, the effect on r is $+0.004$ with a p-value of 0.93; if we include the five major components of net wealth, we have a joint effect that has a p-value of 0.45, and no component has an individual effect with a p-value below 0.23. When we include the components of net wealth and some basic demographics (gender, age, marital status, household size, and net income), we do find a significant joint effect of the components of net wealth with a p-value of 0.005, and the individual component net deposits (with financial institutions) has a significant individual effect of -0.07 with a p-value of 0.003. These results point to the importance of controlling for heterogeneity, and we do that in the online appendix by estimating the model for each individual, thereby allowing implicitly for all observable and unobservable individual characteristics.

C. *Payoff Calibration Implications for EUT*

Using these estimates and the average value of wealth in Denmark, we can evaluate the certainty equivalents (CE) of a range of lotteries varying in the scale of the stakes. Implausible implications for large stakes can be detected through an extremely low ratio of CE to the expected value (EV).[19]

Panel B of table 2 shows implied CE values using the utility function, equation (4), and the parameter estimates in panel A. Let H denote a high prize and L denote a low prize, for $H > L$. The CE in table 2 is then the sure amount of money that has the same expected utility to the individual as the lottery that pays H with probability p and L with probability $(1 - p)$. The CE is defined by

$$U(w, CE) = p \times U(w, H) + (1 - p) \times U(w, L). \qquad (14)$$

So this CE solves for risky income in the experiment, and the stakes are chosen to be within the payoff domain in our experiments. The smallest ratio of CE to EV is 0.362, and most are much higher. These ratios are hardly implausible in the sense of the term used by Hansson (1988), Rabin (2000), Neilson (2001), Rieger and Wang (2006), Cox and Sadiraj (2006), and Safra and Segal (2008).

Figure 2 evaluates the traditional Arrow-Pratt measure of relative risk aversion (RRA) in equation (5) for the estimated EUT-PAI model. The wealth levels in each panel range up to 5 million kroner. Panel A displays RRA for low-stakes lottery prizes up to 10,000 kroner, and panel B displays RRA for high-stakes lottery prizes up to 1 million kroner. Both panels A and B show modest levels of risk aversion for a wide range of wealth and experimental payoffs.

These PAI estimates allow us to verify that (a) getting 190 with probability ½ and 0 with probability ½ is rejected in favor of getting 75 for sure, for all wealth amounts smaller than 35 million, and (b) the same utility function exhibits plausible risk aversion for large stakes. Under FAI, no EUT-consistent agent can exhibit both (a) and (b).

It is, however, possible to come up with some edge cases in which the predictions of EUT-PAI are implausible. For example, at a wealth level of 307 kroner, a low prize of 0, and a high prize of 5,000 kroner, we get very low ratios of CE to EV, between 0.0004 and 0.12, for probabilities between 0.01 and 0.3 on the large prize. As the wealth level increases to the mean wealth level of 3,074,678 kroner, the

---

[19] Similar results are obtained with median wealth instead of average wealth. The ratio of EV to CE is slightly lower, but close to those reported here.

FIGURE 2.—ARROW-PRATT RELATIVE RISK AVERSION FOR ESTIMATED EUT-PAI MODEL

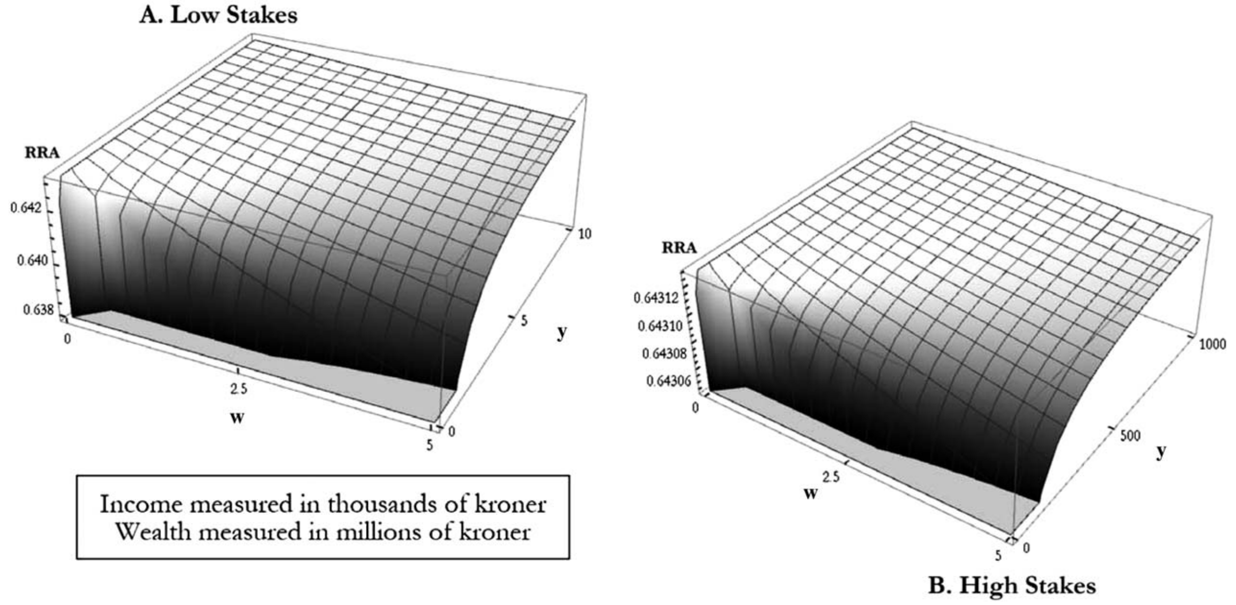


TABLE 3.—ESTIMATES USING RDU-PAI MODEL

A. Estimates

Parameter	Point Estimate	Standard Error	p-value	95% Confidence Interval	
r	0.48	0.05	<0.001	0.38	0.57
η	1.12	0.04	<0.001	1.04	1.2
φ	0.84	0.02	<0.001	0.8	0.88
ω	0.0000106	0.00001	0.46	−0.00002	0.00001
ρ	1	0.00003	<0.001	0.999	1
μ	0.1	0.005	<0.001	0.09	0.11

B. Certainty-Equivalent Calculations with Average Wealth

Large Prize (DKK)	Probability of Large Prize	Small Prize (DKK)	Expected Value (DKK)	Certainty Equivalent (DKK)	Ratio
200	0.5	100	150	141	0.937
500	0.5	100	300	244	0.813
1,000	0.5	100	550	395	0.717
2,000	0.5	100	1,050	668	0.636
5,000	0.5	100	2,550	1,418	0.556
5,000	0.01	100	149	126	0.848
5,000	0.1	100	590	290	0.492
5,000	0.3	100	1,570	751	0.478
5,000	0.7	100	3,530	2,371	0.672
5,000	0.9	100	4,510	3,800	0.842

Sample of 442 individuals making 26,520 choices of strict preference. Log likelihood = −16,973 (−16,976 for NAI and −17,049 for FAI). Null hypothesis for p-value results is that the coefficient estimate is 0.

same example generates low ratios between 0.02 and 0.12 for probabilities between 0.01 and 0.2 on the high prize. We return to compare results for these special cases when we allow for RDU risk preferences.

D. Probability Weighting

The RDU model estimates with the PAI specification in panel A of table 3 show evidence of slight probability weighting pessimism. Compared to the EUT estimates for the PAI specification, there is less curvature on the utility of outcomes once the possibility of probability pessimism is allowed for.[20] We can easily reject the assumption that there is no probability weighting ($\eta = \varphi = 1$), and this is reflected in the improved log likelihood with the RDU model over EUT. In terms of PAI, the estimates are similar to those under EUT except that there is slightly more substi-

---

[20] In other words, for the same choice data, the EUT and RDU models decompose the same risk premium in a different way. The EUT model ascribes all of the risk premium to $U'' < 0$, and the RDU model explains the risk premium with $U'' < 0$ as well as probability pessimism. Since probability pessimism, ceteris paribus U, generates a risk premium itself, the net effect must be for there to be less diminishing marginal utility under RDU than there is under EUT.

tutability between wealth and lab payoffs—in particular, the fundamental finding that $\omega \approx 0$ is the same.

The overall log likelihood of the RDU-PAI model is the best of the RDU specifications considered (RDU-NAI, RDU-PAI, and RDU-FAI). We can formally reject the FAI hypotheses since ω is estimated precisely, $\omega \approx 0$, and we cannot formally reject the null hypothesis that $\omega = 0$ at any conventional statistical level. For the same reasons, we cannot reject the NAI hypothesis either.

For reasons already noted for the EUT-PAI model, when $\omega \rightarrow 0$, the economic meaning of the parameters defining the substitutability of w and y disappears. We formally estimate ρ to be 0.9999927, with a standard error that spans 1, so it is no surprise that the estimate of $\sigma = 1/(1 - \rho)$ is extremely high, at 137,913, and with a large standard error. Again, these wild numerical values follow directly from the economics of the CES function (2) when $\omega \rightarrow 0$, and have no substantive significance or effect on the other parameter estimates (i.e., one could just as easily have constrained $\rho = 1$ and inferred essentially the same estimates).

E. *Payoff Calibration Implications for RDU*

Using the RDU-PAI estimates from table 3, we can again evaluate the ratio of the *CE* to the *EV* for a range of low-stakes and high-stakes lotteries. Using the same lotteries as in panel B of table 2, in panel B of table 3, the *CE* now solves

$$U(w, CE) = d(p) \times U(w, H) + (1 - d(p)) \times U(w, L). \tag{15}$$

The smallest ratio of *CE* to *EV* in table 3 is 0.478, and most are much higher, exactly as in table 2. In general, the ratios in tables 2 and 3 are similar. It is easy to verify that the RDU-PAI model also satisfies the payoff calibration conditions noted earlier for the EUT-PAI model.

Again, as with the EUT-PAI estimates, using these RDU-PAI estimates, one can verify that (a) getting 190 with probability ½ and 0 with probability ½ is rejected in favor of getting 75 for sure, for all wealth amounts smaller than 15.8 million and (b) the same utility function exhibits plausible risk aversion for large stakes. Under FAI, no RDU-consistent agent can exhibit both (a) and (b).[21]

Using these RDU estimates, we can reconsider the edge cases noted earlier, under EUT-PAI, in which the PAI predictions are implausible. Under EUT-PAI, at the low wealth level of 307, the ratio of *CE* to *EV* was between 0.0004 and 0.12 for probabilities between 0.01 and 0.3 on the large prize; with RDU-PAI these ratios are between 0.04 and 0.27, which range from implausible to plausible. The ratio is 0.09, 0.13, and 0.20 for probabilities on the large prize of 0.05, 0.1, and 0.2, respectively. As the wealth level increases to the mean wealth level of 3,074,678 kroner, the same example generates plausible ratios under RDU-PAI between 0.26 and 0.31 for probabilities between 0.01 and 0.2 on the high prize.

These edge cases show that although the PAI model can accommodate the risk version at small and large stakes at the same time, there remain cases falsifying the model. These edge cases allow us to identify the limits of the PAI approach as it is specified here. However, considering a more flexible specification of ω, where it varies with context, could accommodate these edge cases. When RDU-PAI fails to work in these edge cases, so does RDU-NAI. However, the RDU-PAI prediction becomes plausible at wealth levels that are large enough to make baseline wealth $\omega \times w$ meaningful for predictions with stochastic income. In contrast, the performance of RDU-NAI cannot improve with increasing wealth levels. This also applies to cumulative prospect theory, which is equivalent to RDU-NAI when all choices are made on the gain domain. With the exception of the edge cases, the PAI model does well, as illustrated by the examples in tables 2 and 3. It does particularly well when paired with the RDU model of decision making under risk.

VI. Generalizations

As flexible as our approach is in comparison to the full integration and no integration special cases that have dominated the discussion, it is still something of a reduced-form approach to the structural question of the joint determination of lab and nonlab choices. In effect, we take the myriad of decisions underlying w to be given, implicitly assuming that all components of w are symmetric in their relation to y. Given the importance of the issue, we sketch several deeper issues that must be addressed as one generalizes our approach.

In general, it need not be the case that there is symmetry with respect to components of w and experimental choices over y. This is immediately problematic when one considers experimental interventions in the field that offer choices over vectors of commodities rather than just money. For example, the experimental provision of a subsidized micro-insurance product over one type of stochastic outcome, such as the weather, might be expected to interact with cropping choices in a different way from family planning decisions or retirement decisions. Closer to our setting, some components of w, such as more liquid components of wealth,

---

[21] Although these exercises showing how a representative agent would react to various risky contexts are informative about average behavior, they do not allow for heterogeneity in preferences. In fact, the estimate of ω may in part reflect heterogeneity in risk attitudes that just happens to be correlated with wealth rather than some true relation between risk attitudes and wealth. Under power utility A, for any given value of r, a higher wealth level would predict more risk-taking choices in the lottery tasks. Without having observations where wealth varies at the individual subject level, this possibility cannot be ruled out. Thus, if the true preferences are NAI, a positive ω could just be reflecting the possibility that in our sample, the subjects with higher wealth are less risk averse. Or if the true preferences are FAI, $\omega < 1$ could just be reflecting the possibility that in our sample, the subjects with higher wealth are more risk averse.

might be viewed as closer substitutes to experimental income than others.[22] These extensions can be immediately captured with nested-CES aggregator functions of the kind that are common in demand analysis and computable general equilibrium modeling.[23]

In a related vein, individual wealth might be viewed as a closer substitute to experimental income that the individual is choosing over, and other household wealth as not perfectly fungible with individual wealth. Or we might consider an intertemporal utility function defined over stochastic prizes to be paid today and stochastic prizes to be paid in the future (Kihlstrom, 2009; Andersen et al., 2018).[24] In essence, wealth held as financial assets is simply a claim on future income in this manner, thus motivating interest in such intertemporal utility functions.

Once we consider multiple arguments of the utility function, there are a number of theoretical subtleties to consider. One issue is to consider multivariate measures of risk aversion. Kihlstrom and Mirman (1974) proposed such an approach under the restrictive assumption that the ordinal preferences underlying two expected utility functions exhibit the same preferences over nonstochastic outcomes. In this case, they propose a scalar measure of total risk aversion that allows one to make statements about whether one person is more risk averse than another in several dimensions or if the same person is more risk averse after some event than before.

If one relaxes this assumption, which is not an attractive one, Duncan (1977) shows that the Kihlstrom and Mirman (1974) multivariate measure of risk aversion naturally becomes matrix valued. Hence, one has vector-valued risk premiums, and this vector is not direction dependent in terms of evaluation. Karni (1979) shows that one can define the risk premium in terms of the expenditure function rather than the direct utility function, and then evaluate it uniquely by further specifying some statistic of the stochastic process. For example, if one is considering risk attitudes toward a vector of stochastic price shocks, then one could use the mean of those shocks.

A closely related literature defines multiattribute risk aversion where the utility function is defined over more than one attribute. In our case, one attribute would be experimental payoffs y, and the other attribute would be extra experimental wealth w. In this context, Keeney (1973) first defined the concept of conditional risk aversion, Richard (1975) defined the same concept as bivariate risk aversion, and Epstein and Tanny (1980) defined it as correlation aversion. There are several ways to extend these pairwise concepts of risk aversion over two attributes to more than two attributes, as reviewed by Dorfleitner and Krapp (2007).

One attraction of the concept of multiattribute risk aversion is that it allows a relatively simple characterization of the functional forms for utility that rule out multiattribute risk attitudes: additivity. One can have an additive multiattribute utility function and still exhibit partial, or single-attribute, risk aversion. Similarly, one can generate results that do not depend on partial, single-attribute risk aversion but could still depend on multiattribute risk aversion.[25]

A simple but important application of the concept of multiattribute aversion, referred to above as correlation aversion, is when considering intertemporal utility functions. In this case, allowing for a nonadditive intertemporal utility function allows one to tease apart atemporal risk preferences from time preferences, especially temporally correlated risk preferences. In this application, one attribute is the amount of money involved (more or less), and the other attribute is when it is paid (sooner or later). This approach can be directly implemented in controlled experiments, as illustrated by Andersen et al. (2016). For present purposes, it can be viewed as another application of the idea of bivariate risk aversion, which is the same idea as our concept of partial asset integration over atemporal w and y.

A second broad set of issues is the characterization of behavior when portfolio choices are disaggregated and when they are integrated with consumption and leisure choices. Within the field of insurance economics, Mayers and Smith (1983) and Doherty (1984) have stressed the confounding effect that allowing for nontraded assets can have on the demand for insurance. For example, if risks in one domain are perfectly correlated with risks in another domain but traded insurance is available in only one domain, the rational risk-averse agent would tend to "over-insure." A large part of the theory of risk management derives from the complementarity and substitutability of "self-protection" and "self-insurance" activities with formal insurance purchases identified by Ehrlich and Becker (1972). The joint modeling of consumption behavior, leisure demand, and portfolio choices begun with nonadditive

---

[22] We can consider those subjects who have more than the median fraction of net wealth in relatively liquid form, which in our case refers to net assets in financial institutions, bonds, and shares. For simplicity of interpretation, we focus just on point estimates for individual subjects, without conditioning on the statistical significance of the estimate. Around 77% of these subjects are RDU consistent. Just over 92% of these subjects have an ω less than 0.05, and 85% have an ω less than 0.001; 79%, 83%, and 90%, respectively, have a weighted baseline wealth $\omega \times w$ less than 10 kroner, 1,000 kroner, and 100,000 kroner, respectively. Just over 86% of these subjects have a coefficient of relative risk aversion for the composite, r, greater than 0 and less than 0.5. Hence we conclude that these subjects are actually closer to NAI than the typical subject.

[23] The nested CES class allows global regularity and local flexibility in the specification proposed by Perroni and Rutherford (1995). Many specifications that allow local flexibility trade off global regularity, an important property for calibration critiques.

[24] One might argue that some of these examples of imperfect substitutes derive from the absence of perfect capital markets. For example, in the intertemporal case, the existence of perfect capital markets implies the familiar Fisherian (non-)separation theorem. In these cases, one would simply restate results in terms of indirect utility functions.

[25] For multivariate risk aversion, the Hessian should be negative semidefinite under the Kihlstrom and Mirman (1974) definition. For positive r, our utility function, equation (4), is a composition of increasing, concave functions; hence, its Hessian is negative semidefinite. Applying the matrix-valued measures of Duncan (1977) and Karni (1979) would be more involved, of course.

utility functions by Cox (1975) and Ingersoll (1992) identifies numerous avenues for testable propositions about the unexpected spillover effects of policy interventions. There is also a large literature on the effects of consumption commitments on behavior toward risk, starting with Grossman and Laroque (1990) and applied directly to the issue of risk calibration by Chetty and Szeidl (2007). Finally, the partial asset integration approach could provide a rigorous bridge to characterizing the manner in which decision makers employ mental accounts to structure the trade-offs between components of w and y, in the spirit of Thaler (1985) and Thaler and Johnson (1990).[26] The hypothesis of mental accounts involves testable statements about the nested nature of substitutability between different components of w and/or y, and the possibility that ω is context dependent. Once we consider a wider range of stakes for both income and wealth, there are many ways of characterizing the relationship between risk attitudes over these utility arguments. Such specifications are discussed in the broader literature on multivariate and multiattribute risk aversion.

A third set of broad issues has to do with the treatment of wealth as being deterministic and known, while experimental income is stochastic by experimental design. Although consistent with the manner in which asset integration is discussed in the literature, our PAI approach formally allows for there to be a joint probability distribution over wealth and experimental income. An important extension would be to elicit subjective beliefs from individuals about the value of their net wealth at the time of the experiment (or as of some very recent date). After all, who knows with certainty the current value of their net wealth? Since the correlation between subjective beliefs about own wealth and experimental income is 0, again by design, one can just elicit beliefs about wealth (Harrison et al., 2017) and then construct the joint distribution as a mixture of subjective beliefs about own wealth and objective probabilities in the experimental lotteries.

This extension connects our approach to the logic of Barberis et al. (2006), who emphasize the role of risks from gambles such as one confronts in an experiment being merged with preexisting risks from extraexperimental income or wealth. If the risks in the experimental lottery are independent of these preexisting risks, the diversification benefits of the combination might offset any first-order risk aversion toward the experimental lottery evaluated in isolation. Barberis et al. (2006) then posit that the individual evaluates small-stakes gambles in isolation and is driven to exhibit first-order risk aversion, but that the same agent evaluates large-stakes gambles as part of this broader portfolio, tempering the small-stakes risk aversion. Our approach does not require this state-dependent utility specification to account for small-stakes risks and large-stakes risks, although we certainly agree that the riskiness of wealth and experimental income ought to be considered jointly in a complete treatment.

This extension also connects our approach to the logic of Köszegi and Rabin (2007), who consider the implications of loss aversion relative to a stochastic reference point, defined in terms of subjective beliefs about outcomes of the lottery. Recognizing that "relatively little evidence on the determinants of reference points currently exists" (p. 1051), they make this notion operational by assuming that individuals use the EV of the lottery as their subjective belief about the lottery outcome. Our approach immediately extends to include this specification, since we formally allow a joint probability distribution over wealth and experimental income.

The theme of these comments is that our approach is much more general than the resolution of a puzzle about the calibration of choices over risky y in the lab when one takes into account extra lab w. In effect, the rigorous evaluation of seemingly arcane calibration puzzles via models of partial asset integration opens up many areas for research that have tended to be neglected in the calibration debate.

VII. Conclusions

The experimental behavior of adult Danes who have any personal wealth is consistent with partial asset integration, in the dual sense that they behave as if some fraction of personal wealth is combined with experimental prizes in a utility function and that the combination entails less than perfect substitution. Of course, those who have no wealth cannot, as a matter of definition, integrate it with experimental income. Overall, we conclude that our subjects do not perfectly asset integrate.

The implied risk attitudes from estimating these partial asset integration specifications imply risk premiums and certainty equivalents under EUT that are a priori plausible when confronted with the payoff calibration paradox.

---

[26] Thaler and Johnson (1990) focused directly on the question of how risk-taking behavior is affected by prior gains or losses and do not directly consider integration with wealth. But the issues they examine with respect to the components of y have direct application to the generalization we propose. They view choices from the perspective of prospect theory (PT) but allow for interesting variations in the manner in which the editing phase of PT is applied. They provide a simple example in which the subject is told that he or she has just won \$30 and must then choose between (a) no further gain or loss or (b) a fifty-fifty chance of winning \$9 or losing \$9. Three representations of this problem are suggested: (a) $u(\$21) + w(½)\,[u(\$39) - u(\$21)]$; (b) $u(\$30) + w(½)\,u(\$9) + w(½)\,u(-\$9)$; and (c) $u(\$21) + w(½)\,u(\$18)$. The representation in (a) assumes that prior outcomes are embedded into the choice problem. In effect, it adds memory to the standard PT representation of the task and then applies the PT editing rule that the prospect is broken into the certain part and then the residual uncertain part (Kahneman & Tversky, 1979). The representation in (b) assumes that prior outcomes, in this case the \$30 of cumulative income, has no effect on the framing of the task. This is the standard PT formulation. The difference between (a) and (b) has something of the flavor of the asset integration parameter ω that we introduced. But it also has something of the flavor of an endogenous reference point for PT. The representation in (c) assumes that subjects actively deform the prospect to make it appear more attractive. Thus, the possibility of a \$9 loss is integrated into the \$30 on hand, to be evaluated as a certain \$21, and the risky part of the gamble is evaluated as a potential gain of \$18.

Hence, our EUT-PAI specification is promising by surviving the payoff calibration paradox.

Extending the analysis to an RDU model, we find evidence of modest probability weighting and diminishing marginal utility under partial asset integration. Only when one insists a priori, and contrary to the inferences we draw about behavior, that decisions are best characterized with full asset integration does probability weighting come to dominate the characterization of risk attitudes over experimental payoffs. Nonetheless, the RDU-PAI specification also seems to survive the payoff calibration paradox.

These are constructive solutions to the payoff calibration paradoxes. In addition, the rigorous, structural modeling of partial asset integration points to a rich array of neglected questions in risk management and policy evaluation in important field settings.

REFERENCES

Andersen, Steffen, Glenn W. Harrison, Morten I. Lau, and E. Elisabet Rutström, "Discounting Behavior: A Reconsideration," *European Economic Review* 71 (2014), 15–33.

——— "Multiattribute Utility, Intertemporal Utility and Correlation Aversion," *International Economic Review* 59 (2018), 537–555.

Arrow, Kenneth J., "The Role of Securities in the Optimal Allocation of Risk-Bearing," *Review of Economic Studies* 31 (1964), 91–96.

——— *Essays in the Theory of Risk-Bearing* (Chicago: Markham, 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* 96 (2006), 1069–1090.

Chetty, Raj, and Adam Szeidl, "Consumption Commitments and Risk Preferences," *Quarterly Journal of Economics* 122 (2007), 831–877.

Cox, James C., "Portfolio Choice and Saving in an Optimal Consumption-Leisure Plan," *Review of Economic Studies* 42 (1975), 105–116.

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

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

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

Cox, James C., Vjollca Sadiraj, Bodo Vogt, and Utteeyo Dasgupta, "Is There a Plausible Theory for Decision under Risk? A Dual Calibration Critique," *Economic Theory* 54 (2013), 305–333.

Debreu, Gerard, *Theory of Value* (New York: Wiley, 1959).

——— "Least Concave Utility Functions," *Journal of Mathematical Economics* 3 (1976), 121–129.

Doherty, Neil A., "Portfolio Efficient Insurance Buying Strategies," *Journal of Risk and Insurance* 51 (1984), 205–224.

Dorfleitner, Gregor, and Michael Krapp, "On Multiatttributive Risk Aversion: Some Clarifying Results," *Review of Managerial Science* 1 (2007), 47–63.

Duncan, George T., "A Matrix Measure of Multivariate Local Risk Aversion," *Econometrica* 45 (1977), 895–903.

Ehrlich, Isaac, and Gary S. Becker, "Market Insurance, Self-Insurance, and Self-Protection," *Journal of Political Economy* 80 (1972), 623–648.

Epstein, Larry G., and Stephen M. Tanny, "Increasing Generalized Correlation: A Definition and Some Economic Consequences," *Canadian Journal of Economics* 13 (1980), 16–34.

Grossman, Sanford J., and Guy Laroque, "Asset Pricing and Optimal Portfolio Choice in the Presence of Illiquid Durable Consumption Goods," *Econometrica* 58 (1990), 25–51.

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

Harrison, Glenn W., Morten I. Lau, Don Ross, and J. Todd Swarthout, "Small Stakes Risk Aversion in Experiments: A Reconsideration," *Economics Letters* 160 (2017), 24–28.

Harrison, Glenn W., Jimmy Martínez-Correa, J. Todd Swarthout, and Eric Ulm, "Scoring Rules for Subjective Probability Distributions," *Journal of Economic Behavior & Organization* 134 (2017), 430–448.

Harrison, Glenn W., and E. Elisabet Rutström, "Risk Aversion in the Laboratory," in James C. Cox and Glenn W. Harrison, eds., *Risk Aversion in Experiments* (Bingley, UK: Emerald 2008).

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

Hirshleifer, Jack, "Investment Decision under Uncertainty: Choice-Theoretic Approaches," *Quarterly Journal of Economics* 79 (1965), 509–536.

Ingersoll, Jonathan, "Optimal Consumption and Portfolio Rules with Intertemporally Dependent Utility of Consumption," *Journal of Economic Dynamics and Control* 16 (1992), 681–712.

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

Karni, Edi, "On Multivariate Risk Aversion," *Econometrica* 47 (1979), 1391–1401.

Keeney, Ralph L., "Risk Independence and Multiattributed Utility Functions," *Econometrica* 41 (1973), 27–34.

Kihlstrom, Richard E., "Risk Aversion and the Elasticity of Substitution in General Dynamic Portfolio Theory: Consistent Planning by Forward Looking, Expected Utility Maximizing Investors," *Journal of Mathematical Economics* 45 (2009), 634–663

Kihlstrom, Richard E., and Leonard J. Mirman, "Risk Aversion with Many Commodities," *Journal of Economic Theory* 8 (1974), 361–388.

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

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

Luce, R. Duncan, and Howard Raiffa, *Games and Decisions: Introduction and Critical Survey* (New York: Wiley, 1957).

Mayers, David, and Clifford W. Smith, "The Interdependence of Individual Portfolio Decisions and the Demand for Insurance," *Journal of Political Economy* 91 (1983), 304–311.

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

Perroni, Carlo, and Thomas F. Rutherford, "Regular Flexibility of Nested CES Functions," *European Economic Review* 39 (1995), 335–343.

Pratt, John W., "Risk Aversion in the Small and in the Large," *Econometrica* 32 (1964), 123–136.

Prelec, Drazen, "The Probability Weighting Function," *Econometrica* 66 (1998), 497–527.

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

Richard, Scott F., "Multivariate Risk Aversion, Utility Independence and Separable Utility Functions," *Management Science* 22 (1975), 12–21.

Rieger, Marc Oliver, and Mei Wang, "Cumulative Prospect Theory and the St. Petersburg Paradox," *Economic Theory* 28 (2006), 665–679.

Sadiraj, Vjollca, "Probabilistic Risk Attitudes and Local Risk Aversion: A Paradox," *Theory and Decision* 77 (2014), 443–454.

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

Schechter, Laura, "Risk Aversion and Expected-Utility Theory: A Calibration Exercise," *Journal of Risk and Uncertainty* 35 (2007), 67–76.

Thaler, Richard H., "Mental Accounting and Consumer Choice," *Marketing Science* 4 (1985), 199–214.

Thaler, Richard H., and Eric J. Johnson, “Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice,” *Management Science* 36 (1990), 643–660.

Vickrey, William, “Measuring Marginal Utility by Reactions to Risk,” *Econometrica* 13 (1945), 319–333.

von Neumann, John, and Oskar Morganstern, *Theory of Games and Economic Behavior* (Princeton: Princeton University Press, 1944).

——— *Theory of Games and Economic Behavior*, 3rd ed. (Princeton: Princeton University Press, 1953).

Wakker, Peter P., Ido Erev, and Elke U. Weber, “Comonotonic Independence: The Critical Test between Classical and Rank-Dependent Utility Theories,” *Journal of Risk and Uncertainty* 9 (1994), 195–230.

Wilcox, Nathaniel T., “‘Stochastically More Risk Averse’: A Contextual Theory of Stochastic Discrete Choice under Risk,” *Journal of Econometrics* 162 (2011), 89–104.

——— “Unusual Estimates of Probability Weighting Functions,” Economic Science Institute, Chapman University, working paper (2015).