Estimating Risk Attitudes in Denmark:

A Field Experiment

by

Glenn W. Harrison, Morten I. Lau and E. Elisabet Rutström *

December 2004

Abstract. We estimate individual risk attitudes using controlled experiments in the field in Denmark. These risk preferences are elicited by means of field experiments involving real monetary rewards. The experiments were carried out across Denmark using a representative sample of 253 people between 19 and 75 years of age. Risk attitudes are estimated for various individuals differentiated by socio-demographic characteristics such as income and age. Our results indicate that the average Dane is risk averse, and that risk neutrality is an inappropriate assumption to apply. We also find that risk attitudes vary significantly with respect to several important socio-demographic variables. Our results consistently support the need to recognize the heterogeneity of risk attitudes across individual subjects, as well as the need to use flexible utility functions that do not impose strong restrictions on estimates of risk attitudes on an *a priori* basis. These findings have important implications for the characterization of risk attitudes in policy applications, theoretical modeling, and experimental economics.

* Department of Economics, College of Business Administration, University of Central Florida, USA (Harrison and Rutström), and Centre for Economic and Business Research, Copenhagen, Denmark (Lau). E-mail contacts: GHARRISON@BUS.UCF.EDU, MOL@CEBR.DK, and ERUTSTROM@BUS.UCF.EDU. Rutström thanks the U.S. National Science Foundation for research support under grants NSF/IIS 9817518, NSF/MRI 9871019 and NSF/POWRE 9973669, and Harrison and Lau thank the Danish Social Science Research Council for research support under project #24-02-0124. Steffen Andersen provided superb research assistance and comments throughout. We are grateful to Christian Gollier, John Hey, Chris Starmer, Melonie Sullivan and Peter Wakker for comments. Supporting data, statistical code and experimental instructions are stored in the ExLab Digital Archive at http://exlab.bus.ucf.edu.

Welfare analysis of major economic policy decisions such as those concerning education, employment, health care, international trade and retirement typically assume that individuals affected by the policy are risk neutral. Under such an assumption the expected value of the outcome serves as the measure of an individual's or a group of individuals' expected utility. This assumption is usually made out of convenience, since the true risk attitudes are rarely known. Nevertheless, the judgment that welfare assessments may be seriously biased if individuals are not risk-neutral is not controversial. Since risk attitudes are reflections of subjective preferences, one would expect *a priori* that many would be risk averse.[1]

We elicit measures of individual risk attitudes from a representative sample of the Danish population in order to test three substantive hypotheses. The first hypothesis is that *risk attitudes differ significantly from risk neutrality*, such that the implicit assumption in cost-benefit analysis should be reviewed.[2] The second hypothesis is that there are *identifiable segments of the population across which risk attitudes differ in a systematic way*, such that analysts should allow for observable heterogeneity in the data analysis. The third hypothesis is that *relative risk aversion is not constant* with respect to the income levels of the lottery prizes considered, such that one should avoid popular constant relative risk aversion specifications for policies defined over small income changes.

We use choices with real monetary rewards to elicit risk attitudes and demonstrate the methodological complementarity between lab and field experiments. The choices are based on those used by Holt and Laury [2002], who elicited risk attitudes for university students using

[1] We elicit risk attitudes for individuals. To the extent that the characteristics of individuals are used to define "representative households," we can refer to the individual and the household interchangeably. However, we remain agnostic concerning the way in which the risk attitudes of individual household members are aggregated into one household risk attitude.

[2] Following Rabin [2000], there are some specifications of expected utility theory for which a finding of risk aversion at these levels of income is incoherent. This argument does not apply if expected utility theory is defined over income earned during the experiment, rather than over terminal lifetime wealth. Appendix A reviews this argument, and its relevance for experimental studies of risk aversion.

controlled laboratory experiments. We apply extended versions of these experimental procedures from the lab, but employ subjects that are more representative of individuals affected by public policy changes.[3] Our field experiments were carried out across Denmark for the Danish government, using a nationally representative sample of 253 people between 19 and 75 years of age.

Our results show that *the average Dane is risk averse, and that risk neutrality is an inappropriate assumption to apply*. This finding confirms those reported in Holt and Laury [2002] and Harrison, Johnson, McInnes and Rutström [2003] for American college students. We also find that *risk attitudes do vary significantly with several important socio-demographic variables*. The power of performing a field experiment such as the one reported here is that one can get much greater variation in individual characteristics than is generally found on a college campus. In particular, we find that age affects risk attitudes, and that those 40 and over differ significantly from those who are younger. This is an effect that would simply not be observable using the subjects typically recruited on college campuses, since these subject pools rarely include these older age groups.

In general, relative risk aversion does not vary with the lottery stakes considered here. The *assumption of constant relative risk aversion (CRRA) is therefore acceptable over the domain of income considered if applied to the population as a whole*. Nevertheless, we find significant differences across several demographic characteristics, such that *CRRA is not always appropriate for certain identifiable population sub-groups*. For example, participants who are in the age group 40-49 have a higher RRA for the lowest lottery stakes, but this decreases as the stakes go up. Our findings indicate that much of the variation in demographic effects across various studies may be reconciled by applying a utility function that is more flexible than CRRA, coupled with an explicit allowance for individual heterogeneity.

The implications of our finding that most Danes are risk averse can be significant for welfare

[3] Our experiments are "artefactual field experiments" in the terminology of Harrison and List [2004].

analyses, as shown by a simple back-of-the-envelope calculation. Assume a policy that is predicted to result in either a zero or a positive (2,000 DKK, for example) effect on the income of the average Danish household, with probabilities ¼ and ¾, respectively. The certainty equivalent of this policy, assuming a risk coefficient in the neighborhood of those that we report, is about 40% lower than the expected value of 1,500 DKK. Even if we assumed a 90% confidence in the policy prediction instead of 75% confidence, we would be off by almost 20% if we maintained risk neutrality. In general, it is well recognized that the predicted impacts of large scale policy changes are uncertain[4]. A proper identification of the risk attitudes of affected households can therefore be critical to the accuracy of welfare analysis for policy.

At a methodological level, we demonstrate that it is possible to elicit risk attitudes in a field experiment that reflects the population of a country. We concede that Denmark is a remarkable country in which to recruit subjects and undertake field experiments – over 94% of the field subjects we recruited actually turned up for their sessions! The potential importance of eliciting risk attitudes for policy evaluation justifies the development of procedures to rigorously elicit risk attitudes as one component of large-scale surveys that are routinely conducted in many countries. Our procedures should serve as a "best-case" guide to such efforts in the future.[5]

In section 1 we review our experimental design. We propose several extensions of the basic

[4] In the field of computable general equilibrium models there has long been a recognition that systematic sensitivity analysis of simulations conditioned on uncertain parameters implies uncertain policy impacts: see Bernheim, Scholz and Shoven [1991] and Harrison and Vinod [1992], for example. We provide an extended policy illustration in section 4.

[5] In the Epilogue to a book-length review of the economics of risk and time, Gollier [2001; p.424ff.] writes that "It is quite surprising and disappointing to me that almost 40 years after the establishment of the concept of risk aversion by Pratt and Arrow, our profession has not yet been able to attain a consensus about the measurement of risk aversion. Without such a consensus, there is no hope to quantify optimal portfolios, efficient public risk prevention policies, optimal insurance deductibles, and so on. It is vital that we put more effort on research aimed at refining our knowledge about risk aversion. For unclear reasons, this line of research is not in fashion these days, and it is a shame." He also has similar remarks (pp. 425/6) about the long-standing need for empirical evaluations of restrictive functional forms such as CRRA.

laboratory procedure designed to elicit more precise responses and check for robustness to framing effects. These extensions provide several methodological improvements in the risk elicitation procedure, which are of independent interest. Section 2 explains the field experiments conducted, with additional details on procedures provided in Harrison, Lau, Rutström and Sullivan [2005]. Section 3 examines the results and relates them to those found in the existing literature. We also demonstrate how our design allows one to identify possible framing effects, and evaluate relatively flexible functional forms for risk preferences. Section 4 provides a brief but realistic policy application of the risk attitudes we estimate, to demonstrate the policy importance of allowing for the uncertainty of estimates of welfare impacts on individual households.

1. Experimental Design

A. The Basic Elicitation Procedure

We employ a simple experimental measure for risk aversion called a multiple price list (MPL), previously used by Holt and Laury [2002] (HL) and Harrison, Johnson, McInnes and Rutstrom [2003] (HJMR).[6] Each subject is presented with a choice between two lotteries, which we call A or B. Table 1 illustrates the basic payoff matrix presented to subjects. The first row shows that lottery A offered a 10% chance of receiving 2000DKK and a 90% chance of receiving 1600DKK. The expected value of this lottery, EV^A, is shown in the third-last column as 1640DKK, although the EV columns were not presented to subjects. Similarly, lottery B in the first row has chances of payoffs of 3850DKK and 100DKK, for an expected value of 480DKK. Thus the two lotteries have

[6] The MPL appears to have been first used in pricing experiments by Kahneman, Knetsch and Thaler [1990], and has been adopted in recent discount rate experiments by Coller and Williams [1999] and Harrison, Lau and Williams [2002]. It has a longer history in the elicitation of hypothetical valuation responses in "contingent valuation" survey settings, discussed by Mitchell and Carson [1989; p. 100, fn. 14].

a relatively large difference in expected values, in this case 1170DKK.[7] As one proceeds down the matrix, the expected value of both lotteries increases, but the expected value of lottery B becomes greater than the expected value of lottery A.

The subject chooses A or B in each row, and one row is later selected at random for payout for that subject. The logic behind this test for risk aversion is that only risk-loving subjects would take lottery B in the first row, and only risk-averse subjects would take lottery A in the second last row. Assuming local non-satiation, the last row is simply a test that the subject understood the instructions, and has no relevance for risk aversion at all. A risk neutral subject should switch from choosing A to B when the EV of each is about the same, so a risk-neutral subject would choose A for the first four rows and B thereafter.

These data may be analyzed using a variety of statistical models. Each subject made 10 responses. The responses can be reduced to a scalar if one looks at the *lowest* row in Table 1 at which the subject "switched" over from lottery A to lottery B.[8] This reduces the response to a scalar for each subject and task, but a scalar that takes on integer values between 0 and 10. Alternatively, one could study the effects of experimental conditions in terms of the constant relative risk aversion (CRRA) characterization, employing an interval regression model. The CRRA utility is defined as $U(y) = (y^{1-r})/(1-r)$, where r is the CRRA coefficient.[9] The dependent variable in the interval regression model is the CRRA interval that subjects implicitly choose when they switch from lottery

[7] At the time of the experiment the exchange rate was approximately 6.55DKK per U.S. dollar, so this difference translates into almost $180. The exchange rate was 7.45 per Euro, implying a difference of about €160.

[8] Some subjects switched back and forth several times, but the minimum switch point is always well-defined. It turns out not to make much difference how one handles these "multiple switch" subjects. One explanation to such multiple switching behavior is simply indifference. As will be explained below, in this experiment we explicitly include an "indifference" option, and restrict subject choices so that they cannot switch more than once.

[9] With this parameterization, $r = 0$ denotes risk neutral behavior, $r > 0$ denotes risk aversion, and $r < 0$ denotes risk loving. When $r = 1$, $U(m) = \ln(m)$.

A to lottery B. For each row of Table 1, one can calculate the implied bounds on the CRRA coefficient. These intervals are shown in the final column of Table 1. Thus, for example, a subject that made 5 safe choices and then switched to the risky alternatives would have revealed a CRRA interval between 0.15 and 0.41, and a subject that made 7 safe choices would have revealed a CRRA interval between 0.68 and 0.97, and so on.

B. Extensions

We expand this basic design, with some simple modifications to allow a richer characterization of the utility function and the reliability of the elicitation procedure.

Variations in the Income Domain

We want to allow for changes in the value of prizes, so that we have data for the same subject over more than four prizes and can generate better characterizations of their risk attitudes. We therefore undertake four separate risk aversion tasks with each subject, each with different prizes designed so that all 16 prizes span the range of income that we seek to estimate risk aversion over. Ideally, we would have an even span of prizes over that range of income so that we can evaluate the utility function for the individual at different income levels and know that there were some response at or near that level. The four sets of prizes are as follows, in Danish kroner (DKK), with the two prizes for lottery A listed first and the two prizes for lottery B listed next: (A1: 2000, 1600; B1: 3850, 100), (A2: 2250, 1500; B2: 4000, 500), (A3: 2000, 1750; B3: 4000, 150), and (A4: 2500, 1000; B4: 4500, 50). At the time of the experiments, the exchange rate was approximately 6.55 DKK per U.S. dollar, so these prizes range from approximately $7.65 to $687.

This set of prizes generates an array of possible CRRA values. For example, set 1 generates

CRRA intervals at the switch points of -1.71, -0.95, -0.49, -0.14, 0.15, 0.41, 0.68, 0.97 and 1.37, as shown in Table 1. The other sets generate different CRRA intervals, such that all four sets span 36 distinct CRRA values between -1.84 and 2.21, with roughly 60% of the CRRA values reflecting risk aversion.[10] Any scaling of the prizes that is common within a set will preserve the implied CRRA coefficients, so this design could also be used with smaller or larger payoffs.

We ask the subject to respond to all four risk aversion tasks and then randomly decide which one to play out. Budget constraints precluded paying all subjects, so each subject is given a 10% chance of actually receiving the payment associated with his decision. In our statistical analysis we control for "task effects" by adding binary indicator variables. These will capture order effects as well as other task specific effects.[11]

Iterating the MPL

It is possible to extend the MPL to allow more refined elicitation of the true risk attitude, and yet retain the transparency of the incentives of the basic MPL. We do so in the form of a computerized variant on the basic MPL format which we call an Iterative MPL (iMPL).

The basic MPL is the standard format in which the subject sees a fixed array of paired

---

[10] The second set generates CRRA values of -1.45, -0.72, -0.25, 0.13, 0.47, 0.80, 1.16, 1.59 and 2.21; the third set generates values of -1.84, -1.101, -0.52, -0.14, 0.17, 0.46, 0.75, 1.07 and 1.51; and the fourth set generates values of -0.75, -0.32, -0.05, 0.16, 0.34, 0.52, 0.70, 0.91 and 1.20.

[11] Order has been shown to be a quantitatively significant factor in the context of the HL risk aversion design by Harrison, Johnson, McInnes and Rutström [2003]. In that context the order of the task was confounded with a scaling of all lottery prizes, as well as changes from hypothetical to real payoffs. In our design the stakes are confounded with order although they do not increase monotonically across the tasks. In Andersen, Harrison, Lau and Rutström [2004] we report some complementary laboratory experiments using the same design as in the field, precisely to evaluate the nature of the interaction between the order and other task specific effects. These lab results support the view that there are "learning" effects leading to increases in elicited risk aversion in the final task of four that are independent of the specific lotteries used. There are also some positive effects from the second specific lottery pair that are, similarly, independent of the temporal order in which they are presented to subjects.

options and chooses one for each row. It allows subjects to switch back and forth as they like, and has already been used in many experiments. The iMPL format extends this by first asking the subject to simply choose the row at which he wants to first switch from option A to option B, assuming monotonicity of the underlying preferences to automatically fill out the remaining choices. The second extension of the MPL format is to then allow the individual to make choices from refined options within the option last chosen. That is, if someone decides at some stage to switch from option A to option B between probability values of 0.1 and 0.2, the next stage of an iMPL would then prompt the subject to make more choices *within* this interval, to refine the values elicited.[12]

The iMPL uses the same incentive logic as the MPL. After making all responses, the subject has one row from the first table selected at random by the experimenter. In the MPL that is all there is. In the iMPL, that is all there is if the row selected at random by the experimenter is *not* the one that the subject switched at in the first table. If it *is* the row that the subject switched at, another random draw is made to pick a row in the second table that the subject was presented with, and so on.

As the subject iterates in the iMPL the choices become more and more alike, by design. Hence one would expect that greater cognitive effort would be needed to discriminate between them. At some point we expect the subject to express indifference, which we account for in our analysis by only considering the interval over which the subject could (strictly) discriminate. In fact, one possible explanation to why subjects have been observed switching back and forth between

[12] If the subject always chooses A, or indicates indifference for any of the decision rows, there are no additional decisions required and the task is completed. Furthermore, the iterative format has some "smarts" built into it: when the values being elicited drop to some specified perceptive threshold (e.g., a 1-in-100 die throw), the iMPL collapses down to an endogenous number of final rows and the elicitation task stops iterating after those responses are entered.

choices in MPL is that they are indifferent. If so, explicitly including an indifference option, as we do here, may be a cleaner way to capture this behavior.

Framing Effects

A natural concern with the MPL and iMPL is that it might encourage subjects to pick a response in the middle of the table, independent of true valuations. There could be a psychological bias towards the middle, although that is far from obvious *a priori* since there could also be "endpoint" biases at work.

One solution to this concern, which we find unattractive, is to randomize the order of the rows. This is popular in some experimental studies in psychology and economics which elicit discount rates and risk attitudes using the MPL.[13] We find it unattractive for two reasons. First, if there is a purely psychological anchoring effect towards one part of the table such as the middle, this will do nothing but add noise to the responses. Second, the valuation task is fundamentally harder from a cognitive perspective if one shuffles the order of valuations across rows. This harder task may be worthy of study, but is a needless confound for our inferential purposes.

Framing effects can be relatively easily tested for by varying the cardinal scale of the basic MPL table, or by varying the number of intervals within a given cardinal range. If there is an effect on responses, it will be easy to identify statistically and then to control for in the data analysis. We would not be surprised to find framing effects of this kind. They do not necessarily indicate a failure of the traditional economic model, so much as a need to recognize that subjects in a lab setting use all available information to identify a good valuation for a commodity whenever there is any kind of

[13] Kirby and Maraković [1996], Kirby, Petry and Bickel [1999] and Eckel, Johnson and Montmarquette [2005].

uncertainty involved.[14] Thus it is important to be able to estimate the quantitative effect of certain frames and then allow for them in subsequent statistical analysis. We do not claim that any one frame is the correct one: our goal is simply to build in design checks to assess the importance of them for elicited risk attitudes.

We devise a test for framing effects by varying the cardinal scale of the MPL used in the risk aversion task. Two asymmetric frames are developed: the *skewHI* treatment offers initial probabilities of (0.3, 0.5, 0.7, 0.8, 0.9 and 1), while *skewLO* offers initial probabilities of (0.1, 0.2, 0.3, 0.5, 0.7, and 1). This treatment yields 6 decision rows in Level 1 of the iMPL, as opposed to the 10 rows in the symmetric frame.[15] As suggested by the treatment names, *skewLO* (*skewHI*) is intended to skew responses to be lower (higher) probabilities if subjects pick in the middle.

2. Procedures in Denmark

A. Sampling Procedures

The sample for the field experiments was designed to generate a representative sample of the adult Danish population. There were six steps in the construction of the sample,[16] essentially following those employed in Harrison, Lau and Williams [2002]:

- First, a random sample of 25,000 Danes was drawn from the Danish Civil Registration Office in January 2003. Only Danes born between 1927 and 1983 were included, thereby

---

[14] Harrison, Harstad and Rutström [2004].

[15] The design of the skewed frames does interact with the implementation of the iMPL. In the symmetric frame, all intervals are 10 probability points wide, so that a second level is all that is needed to bring subject choices down to precise intervals of 1 probability point. In the skewed frames, however, because the intervals vary in size, a third level is required to bring choices down to this level of precision, and the number of decision rows in Level 3 depends on the width of the interval in Level 1 at which the subject switches.

[16] Further details are provided in Harrison, Lau, Rutström and Sullivan [2005].

restricting the age range of the target population to between 19 and 75. For each person in this random sample we had access to their name, address, county, municipality, birth date, and sex. Due to the absence of names and/or addresses, 28 of these records were discarded.

- Second, we discarded 17 municipalities (including one county) from the population, due to them being located in extraordinarily remote locations. The population represented in these locations amounts to less than 2% of the Danish population, or 493 individuals in our sample of 25,000 from the Civil Registry.
- Third, we assigned each county either 1 session or 2 sessions, in rough proportionality to the population of the county. In total we assigned 20 sessions. Each session consisted of two sub-sessions at the same locale and date, one at 5pm and another at 8pm, and subjects were allowed to choose which sub-session suited them best.
- Fourth, we divided 6 counties into two sub-groups because the distance between some municipalities in the county and the location of the session would be too large. A weighted random draw was made between the two sub-groups and the location selected, where the weights reflect the relative size of the population in September 2002.
- Fifth, we picked the first 30 or 60 randomly sorted records within each county, depending on the number of sessions allocated to that county. This provided a sub-sample of 600.
- Sixth, we mailed invitations to attend a session to the sub-sample of 600, offering each person a choice of times for the session. Response rates were low in some counties, so another 64 invitations were mailed out in these counties to newly drawn subjects. Everyone that gave a positive response was assigned to a session, and our recruited sample was 268.

Attendance at the experimental sessions was extraordinarily high, including 4 persons who did not respond to the letter of invitation but showed up unexpectedly and participated in the experiment. Four persons turned up for their session, but were not able to participate in the experiments.[17] These experiments were conducted in June of 2003, and a total of 253 subjects participated in the experiments.[18] Sample weights for the subjects in the experiment can be constructed using this experimental design, and are used to calculate weighted distributions and averages that better reflect the adult population of Denmark.

B. Conduct of the Sessions

To minimize travel times for subjects, we reserved hotel meeting rooms in convenient locations across Denmark in which to conduct sessions.[19] Because the sessions lasted for two hours, light refreshments were provided. Participants met in groups of no more than 10. To conduct computerized experiments in the field, it was cost-effective to purchase laptop computers and transport them to the meeting sites. Each subject was identified by a unique ID number. For the randomization procedures, two bingo cages were used in each session, one containing 100 balls and the other containing 3 to 11 balls, depending on the number of decision rows in the iMPL used in

[17] The first person suffered from dementia and could not remember the instructions; the second person was a 76 year old woman who was not able to control the mouse and eventually gave up; the third person had just won a world championship in sailing and was too busy with media interviews to stay for two hours; and the fourth person was sent home because they arrived after the instructions had begun and we had already included one unexpected "walk-in" to fill their position.

[18] Certain events might have plausibly triggered some of the no-shows: for example, 3 men did not turn up on June 11, 2003, but that was the night that the Danish national soccer team played a qualifying game for the European championships against Luxembourg that was not scheduled when we picked session dates.

[19] It is possible to undertake experiments over the web with a large sample of subjects drawn from the population. Kapteyn and Teppa [2003] illustrate how one can elicit hypothetical responses to elicit time preferences using a panel of 2,000 Dutch households connected by home computer to surveys. Although not concerned with risk and time preferences directly, Hey [2002] illustrates how one can augment such electronic panel surveys with real experiments. Donkers and van Soest [1999] elicit hypothetical risk and time preferences from pre-existing panels of Dutch households being surveyed for other reasons. Similar exercises with hypothetical surveys include Hartog, Ferre-i-Carbonnell and Jonker [2002] and van Praag and Booij [2003].

different treatments. We found two bingo cages to be the most transparent and convenient way to generate random outcomes in the experiments.

To begin the sessions, subjects were welcomed and reminded that they were to be paid 500 DKK for their participation to cover travel costs as long as they were able to stay for the full two hours required for the experiment. Anyone who was not able to stay for the full two hours was paid 100 DKK and excused from the experiment. The experimenters then asked for a volunteer to inspect and verify the bingo cages and number of bingo balls.

Instructions for the experiment were provided on the computer screens, and subjects read through the instructions while the experimenter read them aloud. The experimenters followed the same script and procedures for each session, documented in Harrison, Lau, Rutström and Sullivan [2005].

The experiment was conducted in four parts. Part I consisted of a questionnaire collecting subjects' socio-demographic characteristics. Specifically, we collected information on age, gender, size of town the subject resided in, type of residence, primary occupation during the last 12 months, highest level of education, household type (viz., marital status and presence of younger or older children), number of people employed in the household, total household income before taxes, whether the subject is a smoker, and the number of cigarettes smoked per day. Part IV consisted of another questionnaire which elicits information on the subject's financial market instruments, and probes the subject for information on their expectations about their future economic conditions and their own future financial position. The questionnaires are rather long, so we chose to divide them across Parts I and IV in order to reduce subject fatigue and boredom. Part II consisted of the four risk aversion tasks, and Part III presented subjects with the six discount rate tasks similar to those developed in Harrison, Lau and Williams [2002]. We will not discuss the discount rate findings here.

The four risk aversion tasks incorporate the incentive structure and assigned frames described earlier. After subjects completed the four tasks, several random outcomes were generated in order to determine subject payments. For all subjects, one of the four tasks was chosen, then one of the decision rows in that task was chosen. For those subjects whose decision at that row led to the second level of the iMPL table, another random draw was required to choose a decision row in the second level, and yet another random draw was required should that decision have led a subject to the third level in the iMPL. To maintain anonymity we performed the draws without announcing to which subjects it would apply. In the case where a subject indicated indifference for the chosen decision row, another random draw determined whether the subject received the results from Lottery A or Lottery B. At this point all subjects knew whether they were playing Lottery A or Lottery B, and another random draw determined whether subjects were to receive the high payment or the low payment. Finally, a 10-sided die was rolled for each subject. Any subject who received a roll of "0" received actual payment according to that final outcome. All payments were made at the end of the experiment.

A significant amount of time was spent training subjects on the iMPL and the randomization procedures in Part II of the experiment. Subjects were given handouts containing examples of two levels of an iMPL that had been filled in. The training exercise explained the logic of the iMPL and verified that subjects were able to correctly fill in an iMPL as shown in the handout. Next, the experimenters illustrated the random procedures necessary to reach a final lottery outcome for each possible choice in the selected decision row in the first level of the iMPL. Finally, a trainer task was conducted in which payments were in the form of candies. The ten-sided die was rolled for each subject, and candies were given to each subject who received a roll of "0."

3. Results

We present results by answering three questions. First, what is the general level of risk aversion in the Danish population, at least over the domain of income considered here? Specifically, is risk neutrality an acceptable hypothesis? Second, do risk attitudes vary with observable demographics? Related to the possible effect of demographics, we can also ask if there are significant effects on elicited risk attitudes from the task frame or task order. Third, how plausible is it to assume that relative risk aversion is constant over the domain of income considered here?

A. Risk Aversion

Figure 1 shows the observed distribution of risk attitudes in our sample, using the raw mid-point of the elicited interval in the *final* iteration stage of the iMPL, assuming a CRRA utility function over the choices defining each interval.[20] This distribution reflects the symmetric menu treatment, which is the appropriate baseline from which to evaluate the asymmetric menu treatments. For this specification of CRRA, a value of 0 denotes risk neutrality, negative values indicate risk-loving, and positive values indicate risk aversion. Thus we see clear evidence of risk aversion: the mean CRRA coefficient is 0.67, weighted to reflect the Danish population. This distribution is consistent with comparable estimates obtained in the United States, using college students and an MPL design, by Holt and Laury [2002] and Harrison, Johnson, McInnes and Rutström [2003]. Very few subjects are risk loving or risk neutral. Risk aversion is by far a better characterization of the risk preferences of the average Dane.

The unconditional data indicates that there is an effect on elicited risk aversion from the

[20] Thus we are only using the CRRA assumption locally in this instance, to define RRA values for which the subject is indifferent at each binary choice. This is far less restrictive than assuming that CRRA characterizes behavior globally over all choices made by the subject.

framing treatments. Figure 2 displays these data in a manner that allows one to easily compare the effects of the treatment. The *skewLO* treatment resulted in an average CRRA of 0.43, and the *skewHI* treatment resulted in an average CRRA of 0.91, each in the direction predicted *a priori*. Subjects do appear to have a bias towards the middle of the table. Nevertheless, this effect does not change our overall conclusion that respondents are risk averse, and that we can reject the hypothesis of risk neutrality. Further comparison of the effect of these treatments on elicited risk aversion requires that we condition on the observed differences in the samples assigned to each treatment. Although subjects were randomly assigned to treatment, our samples are not large enough to be able to draw reliable conclusions solely on the basis of randomization (nor were they designed to).

B. Heterogeneity

In order to assess the importance of demographics on risk attitudes, we applied statistical models that condition on observable characteristics of the subjects and allow for flexibility with respect to functional form. Table 2 provides the definitions of the explanatory variables and summary statistics. It is clear that our data set is quite different from the standard laboratory set using college students, and that it is much more representative of the target population. Since we also use sample weights based on county, age group, and sex, our findings are likely to be broadly policy relevant for Denmark.

The Expo-Power (EP) function was proposed by Saha [1993] as a general functional form for utility, and employed by HL to examine the effects of scale changes in payoffs on risk attitudes. The EP function can be defined as $u(y) = [1-\exp(-\alpha y^{1-r})]/\alpha$, where y is income and α and r are parameters to be estimated. Relative risk aversion (RRA) is then $r + \alpha(1-r)y^{1-r}$. So RRA varies with income if $\alpha \neq 0$. This function nests CRRA (as α tends to 0) and CARA (as r tends to 0).[21]

Maximum likelihood estimates of the EP model can be used to calculate the RRA for different income levels. The likelihood function we use here employs the same function used by Holt and Laury [2002] to evaluate their laboratory data.[22] One important econometric extension of their approach is to allow each parameter, r and α, to be a separate linear function of the task controls and individual characteristics, where we estimate the coefficients on each of these linear functions. We also allow for the responses of the same subject to be correlated, due to unobserved individual effects.

Table 3 displays the results from maximum likelihood estimation of the EP model. We include treatment dummies for the *skewLO* and *skewHI* treatments, as well as for the different risk preference elicitation tasks, using the first task as the reference. This model allows for the deliberate survey design described earlier. In particular, we allow for the fact that subjects in one county were selected independently of subjects in other counties, as well as the possibility of correlation between responses by the same subject.[23]

[21] The CRRA characterization of risk attitudes is popular in theoretical and applied work, no doubt due to it's tractability. For example, there is a relatively elaborate theory of bidding behavior in first-price auctions that relies on such representations of risk attitudes in order to solve for closed-form Bayesian Nash equilibria. Fortunately, even if CRRA is not globally valid over a given income domain, it may still be locally valid of a subset of that domain.

[22] Their likelihood function takes the ratio of the expected utility of the safe option to the sum of the expected utility of both options, where each expected utility is evaluated conditional on candidate values of α and r. Their likelihood specification also allow for a "noise parameter" to capture stochastic errors associated with the choices of subjects. Our implementation of their likelihood specification replicates their results exactly on their laboratory data. Alternative statistical specifications might be expected to lead to different estimates of risk attitudes, although one would not expect radically different estimates.

[23] The use of clustering to allow for "panel effects" from unobserved individual effects is common in the statistical survey literature. Clustering commonly arises in national field surveys from the fact that physically proximate households are often sampled to save time and money, but it can also arise from more homely sampling procedures. For example, Williams [2000; p.645] notes that it could arise from dental studies that "collect data on each tooth surface for each of several teeth from a set of patients" or "repeated measurements or recurrent events observed on the same person." The procedures for allowing for clustering

We first consider the marginal effects of demographics and treatments. We then consider the total effects by stratifying our sample by several of the characteristics, since many demographic characteristics co-vary. For example, the men in our sample have a number of characteristics that differ from the women apart from sex: they tend to be younger, have a higher income, more often live in Copenhagen, and are more likely to be employed, a student, skilled with some post-secondary education, and have higher education.

Marginal Effects

Before we investigate the treatment effects based on our conditional analysis, we will see if, and how, responses vary with the demographics of the respondents.[24]

We find a positive and significant effect of sex on r, but not on α, implying (a) that women are more risk averse than men, but (b) neither women nor men have average risk preferences that change with task income. On average, women have a RRA that is 0.09 higher than men, which is not a particularly large difference even if it is statistically significant.

We also find a significant effect from age on risk attitudes. Individuals below 30 are much less risk averse than others, and those above 40 years of age are substantially more risk averse. Thus we observe a major change in risk attitudes as Danes age. This difference is as much as 0.88 when

---

allow heteroskedasticity between and within clusters, as well as autocorrelation within clusters. They are closely related to the "generalized estimating equations" approach to panel estimation in epidemiology (see Liang and Zeger [1986]), and generalize the "robust standard errors" approach popular in econometrics (see Rogers [1993]).

[24] Since EP has two parameters, the concept of a "marginal effect" can actually be used in three senses: is there an effect of the demographic characteristic on r *given* α, is there an effect on α *given* r, and is there an effect on r and α *together*? Since RRA is given by $r + \alpha(1-r)y^{1-r}$, an effect on r translates directly into an effect on RRA at zero income levels or when α=0. If α>0 (<0) then there is increasing (decreasing) RRA over the domain of income considered, such that an insignificant effect on r by itself might be masking lower (higher) RRA at low income levels and higher (lower) RRA at higher income levels. Hence one also has to look at the joint effect of a given characteristic on r and α.

we compare those under 30 to those over 40 or 50.

Single adults have a much lower RRA on average than others. There is no marginal effect from having children, but since marital status is positively correlated with having children the marginal effect of having children is implicitly captured by the effect from being single.

Those with lower income levels tend to have significantly higher risk aversion, although the difference is only 0.10 in terms of RRA. Those with higher income levels have *lower* risk aversion than those in the middle in terms of income, so the difference from the poorest to the rich households is 0.18 in terms of RRA, which can be significant in terms of policy evaluation.

Quire remarkably, there appears to be no evidence that any of the demographic characteristics is associated with RRA that changes over the income domain considered here.

These results also suggest that there is no significant effect on risk aversion from our experimental treatments, or each of the four tasks.

Total Effects

Now consider the total, rather than marginal, effects of key demographic variables. To do so we simply repeat the statistical analysis shown in Table 3, but with only one demographic characteristic included at a time. In this manner our estimates include all of the demographic characteristics correlated with the characteristic of interest. One example mentioned earlier of how the total effects may differ from the marginal effects is that the men in our sample have a number of characteristics that differ from the women apart from sex: they tend to be younger, have a higher income, more often live in Copenhagen, and are more likely to be employed, a student, skilled with some post-secondary education, and have higher education. Several of these characteristics had a significant marginal effect on risk attitudes, hence it is possible that the joint effect of sex along with the characteristics correlated with it could have a significant effect on risk attitudes.

We find that sex has no total effect on risk aversion, despite the marginal effect indicating that women had a higher RRA[25] by about 0.09. The total effect is a statistical wash, with women having a RRA that has a 95% chance of being ±0.04 of the RRA of men.

The only significant total effects come from income. We calculate that lower income households have an RRA that is 0.13 higher than others on average, and higher income households have an RRA that is 0.09 lower than others on average. These total effects are qualitatively consistent with the marginal effects, and are just slightly larger.

There are two total effects that are notable, even if they are not statistically significant at the usual levels. Students have a much lower RRA than others, by about 0.61. The *p*-value for this difference is 0.16. In this case the total effect is roughly double the marginal effect in size. One reason for this increase in size is that the total effect of being young is to also lower one's aversion to risk, specifically to lower average RRA by 0.15 with a *p*-value of 0.21. This is smaller than the marginal effect of 0.46 from Table 3, and less significant. Thus, since students tend to be younger, the total effect of being a student is an amalgam of the marginal effects of being a student and being young, accounting for the larger total effect of being a student compared to the marginal effect of being a student.

One important implication of these results is that one might have to evaluate welfare policies at a dis-aggregated household level in order to correct identify differences in risk attitudes and hence the correct certainty-equivalent of uncertain policies. For example, if one just compared men and

[25] Strictly speaking, these are statements based on the estimated coefficient on the r parameter of the expo-power utility function. But the baseline estimate of the α parameter is not significantly different from zero, nor are any effects of the single demographic characteristic on α statistically significant. Hence the RRA collapses to r, since one effectively has a CRRA function.

women there would be no difference in risk attitudes: the total effect of sex is effectively zero. But if one compared a young Danish man who lived alone and was a student, you would find a significantly less risk averse individual than the average Dane based on the marginal effects from Table 3. This conclusion is of more significance for policy analysis than might be initially appreciated. It is common to find household dis-aggregations provided on just one dimension: age of household, or income of household, or occupation of household. Indeed, in Denmark it is difficult to obtain official household expenditure surveys in which any of these major characteristics are interacted, due to the small sample sizes of each survey cohort. Hence welfare analyses that are restricted to just one of these dimensions may incorrectly treat the individual households as more representative than they actually are.

4. A Policy Illustration

Our policy motivation for eliciting risk attitudes was to obtain a better characterization of the welfare impacts of microeconomic policy alternatives in Denmark. We illustrate the use of the estimates we have collected using a simulation from a computable general equilibrium model developed for the Danish government to evaluate precisely this type of policy.

The specifics of the model are documented in Harrison, Jensen, Lau and Rutherford [2002]. It is basically a static model of the Danish economy calibrated to data from 1992. The version we use has 27 production sectors, each employing intermediate inputs and primary factors to produce output for domestic and overseas consumption. A government agent raises taxes and pays subsidies, and the focus of our policy simulation is on the level of indirect taxes levied by the Danish government. Apart from a representative government household, which consumes goods reflecting public expenditure patterns in 1992, the model differentiates 7 private household types discussed below. The model is calibrated to a wide array of empirical and *a priori* estimates of elasticities of substitution using nested constant elasticity of substitution specifications for production and utility functions. More elaborate versions of the model exist in which inter-temporal and inter-generational behavior are modeled (e.g., Lau [2000]), but this static version is ideal for our illustrative purposes.

The model represents several different private households, based on the breakdown provided by Statistics Denmark from the national household expenditure survey. These household types are differentiated by one of the following characteristics: total household income, disposable household income, occupation, family type, number of people employed, size of town of residence, and type of residence. We select family type for our analysis. This implies that the simulation model identifies welfare impacts of a counter-factual policy simulation on each of 7 households: singles younger than 45 without children, singles older than 45 without children, households younger than 45 without children, households older than 45 without children, singles with children, households with children and where the oldest child is 6 or under, and households with children and where the oldest child is between 7 and 17. The model generates the welfare impact on each household measured in terms of the equivalent variation in annual income for that household. That is, it calculates the amount of income the household would deem to be equivalent to the policy change, which entails changes in factor prices, commodity prices and expenditure patterns. Thus the policy impact is some number of Danish kroner, such as 1,478 DKK, which represents the welfare gain to the household in income terms.

This welfare gain can be viewed directly as the "prize" in a policy lottery. Since there is some uncertainty about the many parameters used to calibrate realistic simulation models of this kind, there is some uncertainty about the calculation of the welfare impact. If we perturb one or more of the elasticities, for example, the welfare gain might well be 1,896 DKK for the household instead of

1,478 DKK. Using randomized factorial designs for such sensitivity analyses, we can undertake a large number of these perturbations and assign a probability weight to each one (Harrison and Vinod [1992]). Each simulation involves a random draw for each elasticity, but where the value drawn reflects estimates of the empirical distribution of the elasticity.[26] We undertake 1,000 simulations with randomly generated elasticity perturbations, so it is as if the household faces a policy lottery consisting of 1,000 distinct prizes that occur with equal probability 0.001. The prizes, again, are the welfare gains that the model solves for in each such simulation.

Figure 4 illustrates the type of policy lottery that can arise. In this case we consider a policy of making all indirect taxes in Denmark uniform, and at a uniform value that just maintains the real value of government expenditure. Thus we solve for a revenue-neutral reform in which the indirect tax distortions arising from inter-sectoral variation in those taxes are reduced to zero. Each box in Figure 4 represents 1,000 welfare evaluations of the model for each household type. The large dot is the median welfare impact, the rectangle is the interquartile range,[27] and the whiskers represent the range of observed values. Thus we see that the policy represents an uncertain lottery for each household, with some uncertainty about the impacts.

If a policy-maker were to evaluate the expected utility to each household from this policy, he would have to take into account the uncertainty of the estimated outcome and the risk attitudes of the household. The traditional approach in policy analysis is to implicitly assume that households are all risk-neutral and simply report the average welfare impact. But we now know that these

---

[26] For example, if the empirical distribution of the elasticity of substitution is specified to be normal with mean 1.3 and standard deviation 0.4, 95% of the random draws will be within $\pm 1.96 \times 0.4$ of the mean. Thus one would rarely see this elasticity take on values greater than 3 or 4 in the course of these random draws.

[27] Defined by the 25th and 75th percentiles, this range represents 50% of the observations around the median.

households are not risk neutral. In fact, stratifying our raw estimates according these household types, we obtain CRRA estimates of 0.74, 0.65, 0.82, 0.44, 0.52, 0.86 and 0.62, respectively, for each of these households. In each case these are statistically significantly different from risk neutrality.

Using these CRRA risk attitude estimates, it is a simple matter to evaluate the utility of the welfare gain in each simulation, then to calculate the expected utility of the proposed policy, and then calculate the certainty-equivalent welfare gain. Doing so reduces the welfare gain relative to the risk-neutral case, of course, since there is some uncertainty about the impacts. For this illustrative policy, this model, and these estimates of risk attitudes, we find that welfare gains are overstated if one neglects risk aversion by 1.6%, 1.4%, 1.8%, 1.1%, 5.1%, 4.6% and 7.9%, respectively, for each of the households. Thus a policy maker would overstate the welfare gains from the policy if risk attitudes were implicitly ignored.

Tax uniformity is a useful pedagogic example, and a staple in public economics, but one that generates relatively precise estimates of welfare gains in most simulation models of this kind. It is easy to consider realistic policy simulations that would generate much more variation in welfare gain, and hence larger corrections from using the household's risk attitude in policy evaluation. For example, assume instead that indirect taxes in this model were reduced across the board by 25%, and that the government effected lump-sum sidepayments to each household to ensure that no household had less than a 1% welfare gain.[28] In this case there are some plausible elasticity configurations for the model that result in very large welfare gains for some households.[29] In this

[28] The manner in which these sidepayments are computed is explained in Harrison, Jensen, Lau and Rutherford [2002]. It corresponds to a stylized version of the type of political balancing act one often encounters behind the scenes in the design of a public policy such as this.

[29] For example, if the elasticity of demand for a product with a large initial indirect tax is higher than the default elasticity, households can substitute towards that product more readily and enjoy a higher real income for any given factor income.

case ignoring the risk attitudes of the household would result in welfare gains being overstated by much more significant amounts, ranging from 18.5% to 41.8% depending on the household.

These policy applications point to the payoff from our estimates of risk attitudes. They are illustrative in several ways. First, we have to assume that the estimates of CRRA obtained from our experimental tasks defined over the domain of prizes up to 4,500 DKK apply more widely, to the domain of welfare gains shown in Figure 4. Given the evidence for CRRA from our analysis of the expo-power functional form, we are prepared to make that assumption for now. But obviously one would want to elicit risk attitudes over wider prize domains to be confident of this assumption. Second, we only have households in Denmark stratified into 7 different household types, each of which is likely to contain households with widely varying characteristics. The current generation of computable general equilibrium models is employing *individual* household data from expenditure surveys, and effectively merging the rich detail of micro-simulation models with the behavioral coherence of a fully-specified general equilibrium model.[30]

5. Conclusions

We elicit attitudes to risk from a sample of individuals representative of the general Danish population using real economic commitments. We find that risk aversion is by far a better characterization of the risk attitudes of most Danes than is risk neutrality. There are also variations

[30] One of the earliest large-scale models by Piggott and Whalley [1985] incorporated 100 households in a model of the United Kingdom, utilizing every available representative household of the Family Expenditure Survey available to them. Households were differentiated by interactions of income, family structure, and occupation. More recent efforts to include multiple households in such models have been driven by concerns about the impacts of trade reform on poverty in developing countries, since one can only examine those by identifying the poorest households: see Harrison, Rutherford and Tarr [2003] and Harrison, Rutherford, Tarr and Gurgel [2004]. Clearly one would expect risk aversion to be a particularly important factor for households close to or below the absolute poverty line.

in risk attitudes across several identifiable socio-demographic characteristics of the Danish population, implying that welfare evaluations of government policies for those individuals should take these differences into account. We find strong support for a decrease in risk aversion as the age of a person increases, particularly after age 40. We also report that there is some support for constancy of relative risk aversion for the Danish population as a whole, although non-constancy clearly exists for some identifiable segments of the population. Our results consistently support the need to recognize the heterogeneity of risk attitudes across individual subjects, and to use flexible utility specifications that do not constrain estimates of risk attitudes *a priori*. These findings have important implications for the characterization of risk attitudes in policy applications, theoretical modeling, and experimental economics.

Table 1: Typical Payoff Table for Risk Aversion Experiments

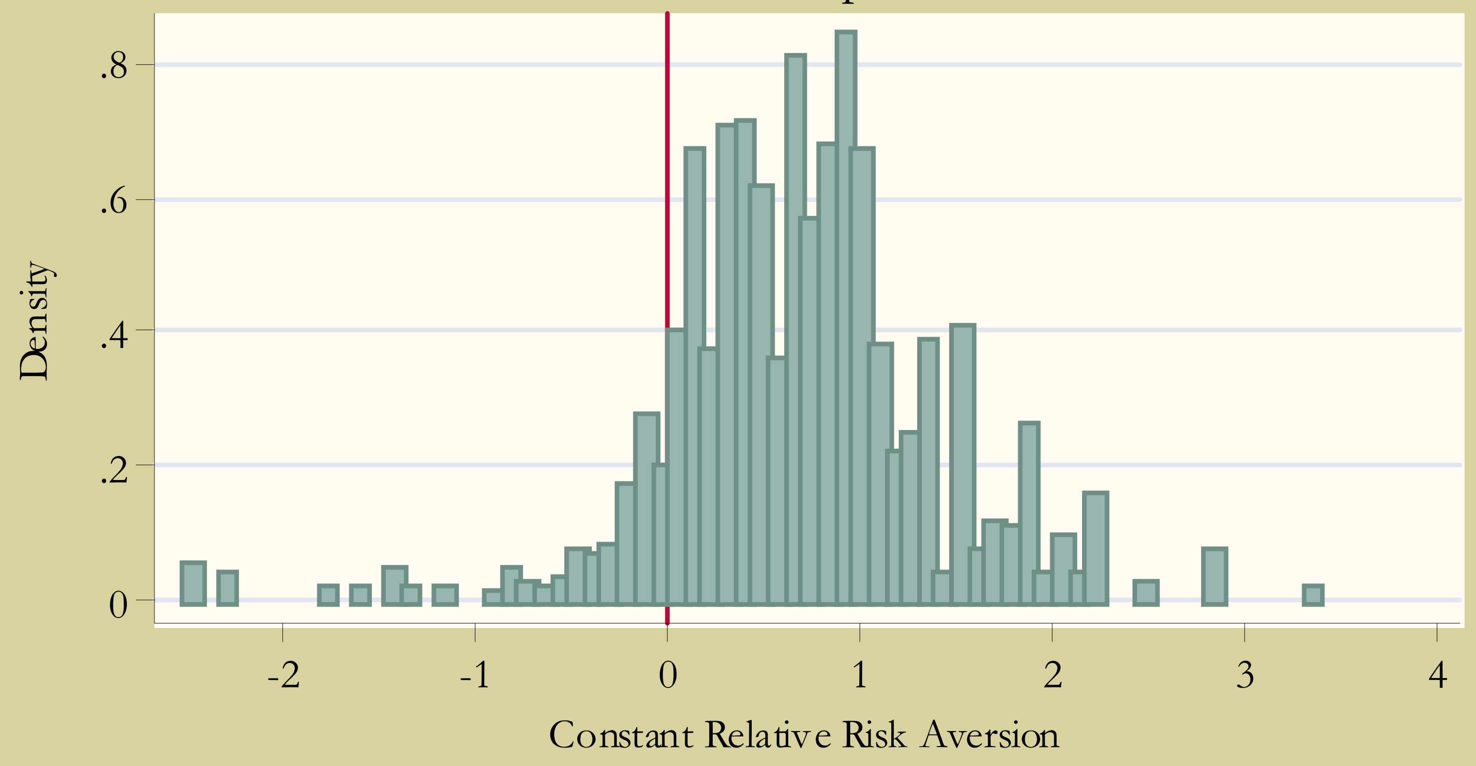
All currency units are Danish kroner (DKK). At the time of the experiment 1 USD = 6.55 DKK = € 7.45.

Lottery A				Lottery B				EV^A	EV^B	Difference	Open CRRA Interval if Subject Switches to Lottery B
p(2000)		p(1600)		p(3850)		p(100)					
0.1	2000	0.9	1600	0.1	3850	0.9	100	1640	480	1170	-∞, -1.71
0.2	2000	0.8	1600	0.2	3850	0.8	100	1680	850	830	-1.71, -0.95
0.3	2000	0.7	1600	0.3	3850	0.7	100	1720	1230	490	-0.95, -0.49
0.4	2000	0.6	1600	0.4	3850	0.6	100	1760	1600	160	-0.49, -0.15
0.5	2000	0.5	1600	0.5	3850	0.5	100	1800	1980	-170	-0.14, 0.14
0.6	2000	0.4	1600	0.6	3850	0.4	100	1840	2350	-510	0.15, 0.41
0.7	2000	0.3	1600	0.7	3850	0.3	100	1880	2730	-840	0.41, 0.68
0.8	2000	0.2	1600	0.8	3850	0.2	100	1920	3100	-1180	0.68, 0.97
0.9	2000	0.1	1600	0.9	3850	0.1	100	1960	3480	-1520	0.97, 1.37
1	2000	0	1600	1	3850	0	100	2000	3850	-1850	1.37, ∞

Note: The last four columns in this table, showing the expected values of the lotteries and the implied CRRA intervals, were not shown to subjects.

Figure 1: Distribution of CRRA in Denmark
With Symmetric Menu

Mid-Point of Raw Responses from iMPL



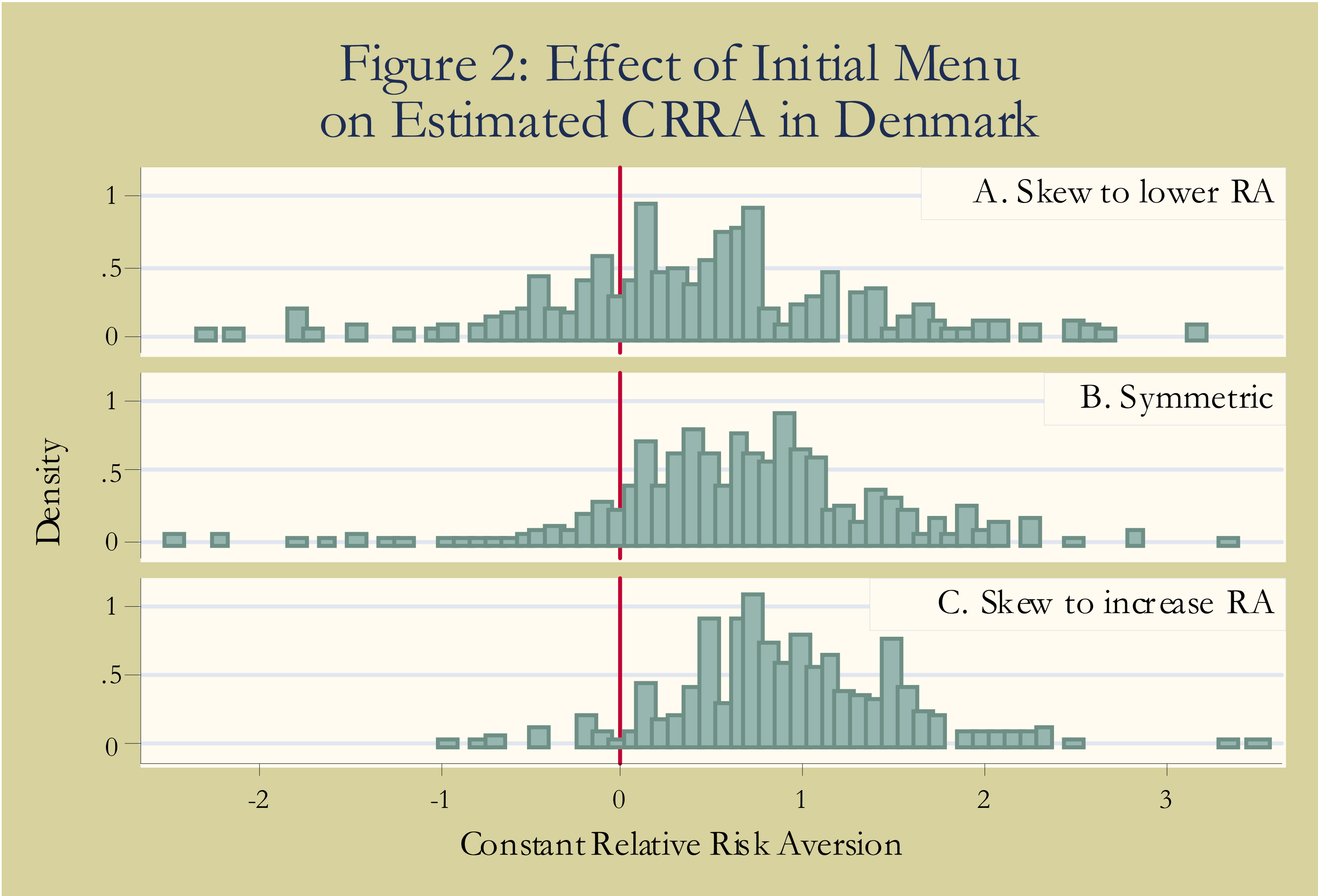


Figure 2: Effect of Initial Menu
on Estimated CRRA in Denmark
A. Skew to lower RA
B. Symmetric
C. Skew to increase RA
Density
1
.5
0
-2
-1
0
1
2
3
Constant Relative Risk Aversion

Table 2: List of Variables and Descriptive Statistics

Variable	Definition	Estimated Population Mean	Raw Sample Mean
female	Female	0.50	0.51
young	Aged less than 30	0.19	0.17
middle	Aged between 40 and 50	0.27	0.28
old	Aged over 50	0.33	0.38
single	Lives alone	0.21	0.20
kids	Has children	0.31	0.28
nhhd	Number of people in the household	2.54	2.49
owner	Owns own home or apartment	0.68	0.69
retired	Retired	0.13	0.16
student	Student	0.10	0.09
skilled	Some post-secondary education	0.38	0.38
longedu	Substantial higher education	0.36	0.36
IncLow	Lower level income	0.33	0.34
IncHigh	Higher level income	0.36	0.34
copen	Lives in greater Copenhagen area	0.27	0.27
city	Lives in larger city of 20,000 or more	0.41	0.39
experimenter	Experimenter Andersen (default is Lau)	0.47	0.49

Legend: Most variables have self-evident definitions. The omitted age group is 30-39. Variable "skilled" indicates if the subject has completed vocational education and training or "short-cycle" higher education, and variable "longedu" indicates the completion of "medium-cycle" higher education or "long-cycle" higher education. These terms for the cycle of education are commonly used by Danes (most short-cycle higher education program last for less than 2 years; medium-cycle higher education lasts 3 to 4 years, and includes training for occupations such as a journalist, primary and lower secondary school teacher, nursery and kindergarten teacher, and ordinary nurse; long-cycle higher education typically lasts 5 years and is offered at Denmark's five ordinary universities, at the business schools and various other institutions such as the Technical University of Denmark, the schools of the Royal Danish Academy of Fine Arts, the Academies of Music, the Schools of Architecture and the Royal Danish School of Pharmacy). Lower incomes are defined in variable "IncLow" by a household income in 2002 below 300,000 kroner. Higher incomes are defined in variable "IncHigh" by a household income of 500,000 kroner or more.

Table 3: Statistical Model of Risk Aversion Responses

Maximum-likelihood estimates of expo-power utility function.
N=15,360 choices by 249 subjects stratified across 12 counties.

Variable	Description	Estimate	Standard Error	p-value	Lower 95% Confidence Interval	Upper 95% Confidence Interval
	A. Estimates of Determinants of r					
Constant		0.40	0.20	0.04	0.01	0.79
skewLO	Skew towards risk loving	-0.01	0.09	0.94	-0.18	0.17
skewHI	Skew towards risk aversion	-0.04	0.10	0.72	-0.24	0.17
Task2	Second risk task	0.02	0.03	0.55	-0.04	0.07
Task3	Third risk task	0.01	0.03	0.72	-0.04	0.06
Task4	Fourth risk task	0.02	0.04	0.65	-0.05	0.09
experimenter	Experimenter effect	0.05	0.04	0.27	-0.04	0.14
female	Female	0.09	0.04	0.01	0.02	0.16
young	Aged less than 30	-0.46	0.11	0.00	-0.67	-0.24
middle	Aged between 40 and 50	0.43	0.15	0.00	0.13	0.74
old	Aged over 50	0.45	0.17	0.01	0.13	0.78
single	Lives alone	-0.17	0.07	0.02	-0.31	-0.03
kids	Has children	-0.01	0.10	0.90	-0.21	0.18
nhhd	Number in household	-0.05	0.04	0.20	-0.12	0.03
owner	Own home or apartment	0.05	0.05	0.36	-0.06	0.15
retired	Retired	-0.07	0.07	0.35	-0.20	0.07
student	Student	-0.30	0.18	0.10	-0.66	0.06
skilled	Some post-secondary education	-0.08	0.06	0.21	-0.20	0.05
longedu	Substantial higher education	0.11	0.07	0.14	-0.04	0.25
IncLow	Lower level income	0.10	0.05	0.04	0.01	0.20
IncHigh	Higher level income	-0.08	0.04	0.06	-0.17	0.00
copen	Lives in Copenhagen area	-0.01	0.07	0.94	-0.15	0.14
	B. Estimates of Determinants of α					
Constant		0.00	0.00	0.29	0.00	0.01
skewLO	Skew towards risk loving	0.00	0.00	0.24	0.00	0.01
skewHI	Skew towards risk aversion	0.00	0.00	0.35	0.00	0.00
Task2	Second risk task	0.00	0.00	0.58	0.00	0.00
Task3	Third risk task	0.00	0.00	0.87	0.00	0.00
Task4	Fourth risk task	0.00	0.00	0.70	0.00	0.00
experimenter	Experimenter effect	0.00	0.00	0.55	0.00	0.00
female	Female	0.00	0.00	0.52	0.00	0.00
young	Aged less than 30	0.00	0.00	0.29	-0.01	0.00
middle	Aged between 40 and 50	-0.02	0.09	0.81	-0.21	0.16
old	Aged over 50	-0.04	0.08	0.60	-0.19	0.11
single	Lives alone	0.00	0.00	0.64	0.00	0.00
kids	Has children	0.00	0.00	0.50	0.00	0.00
nhhd	Number in household	0.00	0.00	0.79	0.00	0.00
owner	Own home or apartment	0.00	0.00	0.45	0.00	0.00
retired	Retired	-0.04	0.10	0.71	-0.24	0.17
student	Student	0.00	0.00	0.40	0.00	0.00
skilled	Some post-secondary education	0.00	0.00	0.86	0.00	0.00
longedu	Substantial higher education	0.00	0.00	0.33	0.00	0.00
IncLow	Lower level income	0.00	0.00	0.33	0.00	0.00
IncHigh	Higher level income	0.00	0.00	0.37	0.00	0.00
copen	Lives in Copenhagen area	0.00	0.00	0.95	0.00	0.00
	C. Estimate of μ					
Constant		0.16	0.02	0.00	0.12	0.20

Note: F-test for the null hypothesis that all coefficients are zero has a value of 7.46 with 21 and 217 degrees of freedom, implying a p-value less than 0.001.

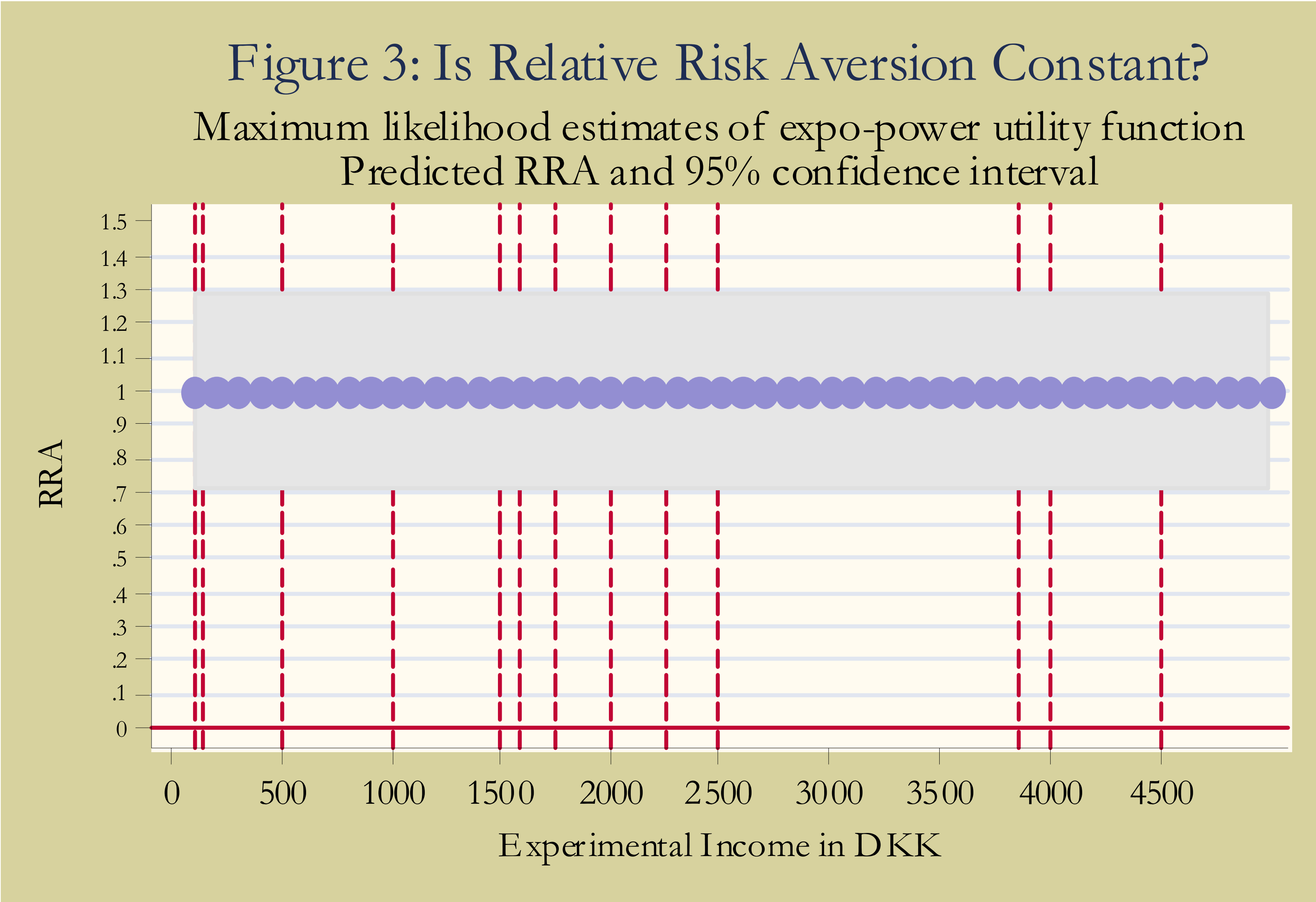
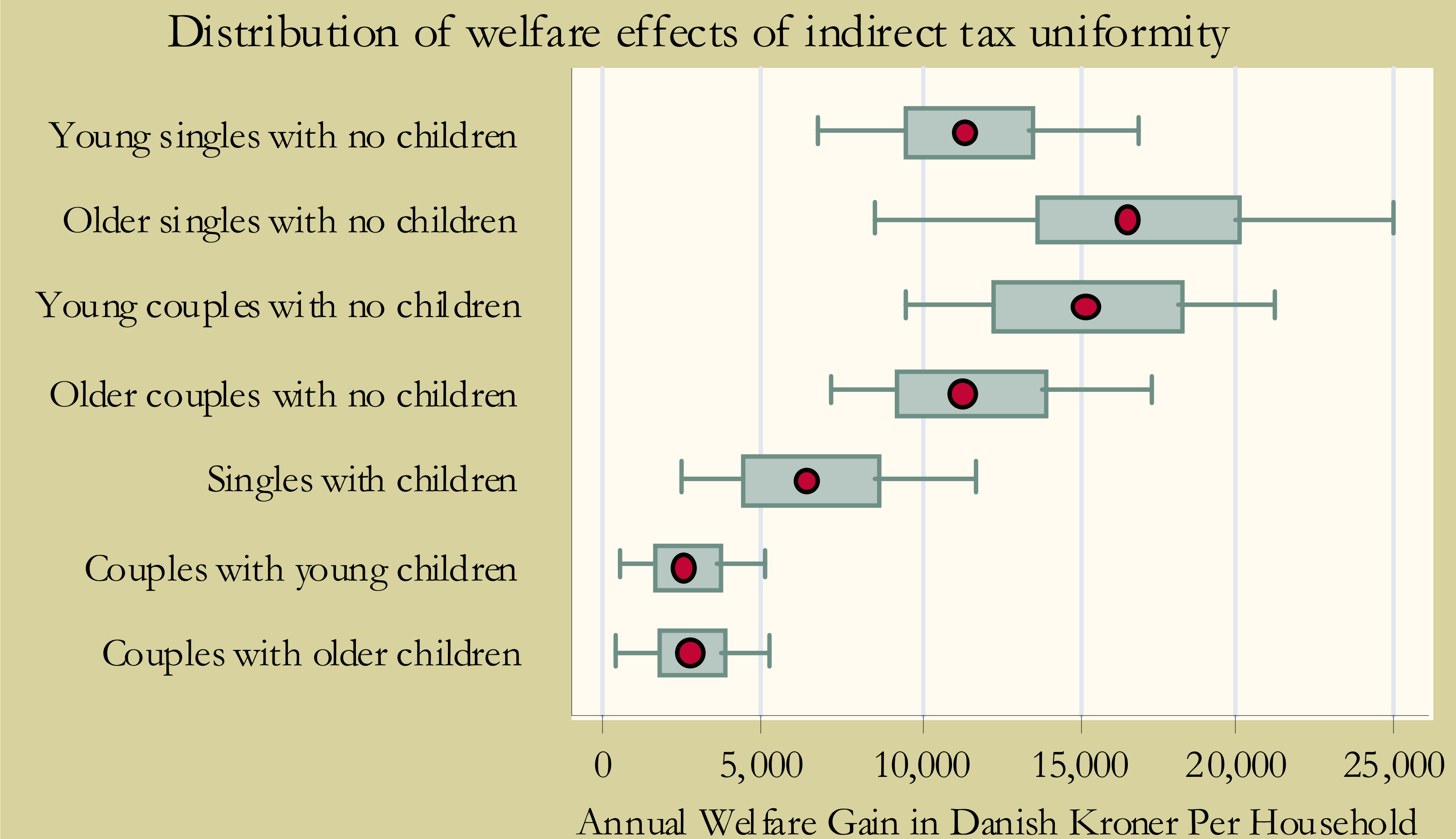


Figure 3: Is Relative Risk Aversion Constant?
Maximum likelihood estimates of expo-power utility function
Predicted RRA and 95% confidence interval
RRA
1.5
1.4
1.3
1.2
1.1
1
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
0
500
1000
1500
2000
2500
3000
3500
4000
4500
Experimental Income in DKK

Figure 4: An Illustrative Policy Lottery



Appendix A: Risk Aversion and Expected Utility Theory

A recent theoretical examination of the role of risk aversion and expected utility theory (EUT) argues that EUT must be rejected for individuals who are risk averse at low monetary stakes. If true, then further tests of EUT are not needed for those individuals who are found to be risk averse in these low stake lottery choices. Rabin [2000] proves a calibration theory showing that if individuals are risk averse over low stakes lotteries then there are absurd implications about the bets those individuals will accept at higher stakes. Rabin [2000] and Rabin and Thaler [2001] allege that this result has general implications for the validity of EUT as a descriptive theory. As explained by Rabin and Thaler [2001; p.222, emphasis added]:

> The logic behind this result is that within the expected utility framework, turning down a moderate stakes gamble means that the marginal utility of money must diminish very quickly. Suppose you have initial wealth of W, and you reject a 50-50 lose \$10/gain \$11 gamble because of diminishing marginal utility of wealth. Then it must be that $U(W + 11) - U(W) \leq U(W) - U(W-10)$. Hence, on average you value each of the dollars between W and W + 11 by at most 10/11 as much as you, on average, value each of the dollars between W-10 and W. By concavity, this implies that you value the dollar W + 11 at most 10/11 as much as you value the dollar W-10. Iterating this observation, if you have the same aversion to the lose \$10/gain \$11 bet at wealth level W + 21, *then you value dollar W + 21 + 11 = W + 32 by at most 10/11 as you value dollar W + 21 - 10 = W + 11, which means you value dollar W + 32 by at most 10/11 × 10/11 ≈ 5/6 as much as dollar W-10*. You will value the W + 210[th] dollar by at most 40 percent as much as dollar W-10, and the W + 900[th] dollar by at most 2 percent as much as dollar W-10. In words, rejecting the 50-50 lose \$10/gain \$11 gamble implies a 10 percent decline in marginal utility for each \$21 in additional lifetime wealth, meaning that the marginal utility plummets for substantial changes in lifetime wealth. You care less than 2 percent as much about an additional dollar when you are \$900 wealthier than you are now. This rate of deterioration for the value of money is absurdly high, and hence leads to absurd risk aversion.

Thus, a problem for EUT does indeed arise if (a) subjects exhibit risk aversion at low stake levels, *and* (b) one assumes that utility is defined in terms of terminal wealth.[31]

---

[31] Terminal wealth refers here to the wealth that the subject has prior to coming into the lab plus any income earned in the lab. Watt [2002] and Palacios-Huerta, Serrano and Volij [2002] argue that the degree of relative risk aversion required for Rabin's result are *a priori* implausible. If an individual turned down a small-stakes gamble with a positive expected return for *any* wealth level, including high wealth levels, then that individual must have extremely high relative risk aversion. Hence it could be reasonable for that individual to turns down more generous gambles at higher stakes.

If, on the other hand, one assumes utility is defined over income, this critique will not apply. Consider the step in the argument that is italicized, and which relies critically on the utility function being defined in terms of terminal wealth. If utility were defined in terms of income, then one could not make this step in the argument: all that one could say would be that the person at wealth level W+21 valued dollar W+11 at most 10/11 as much as he valued the dollar W-10. This is the same statement made in the first step of the argument, so there is no basis for making inferences about how the person values much larger stakes.

A careful reading of Rabin [2000] is consistent with this perspective. Consider this passage (p.1288):

> What *does* explain risk aversion over modest stakes? [...] what is empirically the most firmly established feature of risk preferences, *loss aversion*, is a departure from expected-utility theory that provides a direct explanation for modest-scale risk aversion. Loss aversion says that people are significantly more averse to losses relative to the status quo than they are attracted by gains, and more generally that people's utilities are determined by changes in wealth rather than absolute levels.

One can accept the second contention from the above, that subjects use experimental income (i.e., changes in wealth) rather than absolute levels of wealth as the basis for making decisions, independent of the first point about the asymmetry of risk attitudes either side of the status quo.

Whether or not one models utility as a function of terminal wealth (EUTw) or income (EUTi) depends on the setting. Both specifications have been popular. The EUTw specification was widely employed in the seminal papers defining risk aversion and it's application to portfolio choice. The EUTi specification has been widely employed by auction theorists and experimental economists testing EUT, and it is the specification we employ here.

One is tempted to think that this result is well-known since Markowitz [1952] and Samuelson [1952; ¶13, p.676], but that may just be a hindsight bias. Cox and Sadiraj [2004] and Rubinstein [2002] make these points quite clearly. Cox and Sadiraj [2004] go further to propose a generalization of EUTw and EUTi that allows initial wealth to be an argument of the utility function along with income (as long as initial wealth is not simply added to income, which would be EUTw). They also note that "loss aversion," the alternative favored by Rabin [2000] and Rabin and Thaler

[2001] as a descriptive model of low-stakes risk aversion, is perfectly consistent with EUTi.

Rubinstein [2002] draws the important connection between adopting an EUTi assumption and the question of temporal consistency of preferences, since the income that one received in today's experiment must be "integrated" in some consistent way with the income received in the past (viz., wealth prior to the experiment). This suggests links back to the older literature on the "asset integration hypothesis," reviewed in this context by Quizon, Binswanger and Machina [1984]. In other words, just because one adopts an EUTi characterization and thereby avoid the problems posed by Rabin [2000], one is not free to make any arbitrary assumptions about behavior over time. The laboratory evidence on this matter has it's own controversies: see Frederick, Loewenstein and O'Donoghue [2002] and Coller, Harrison and Rutström [2003].

References

Andersen, Steffen; Harrison, Glenn W.; Lau, Morten Igel, and Rutström, E. Elisabet, "Elicitation Using Multiple Price Lists," *Working Paper 04-08*, Department of Economics, College of Business Administration, University of Central Florida, 2004.

Bernheim, B. Douglas; Scholz, John Karl, and Shoven, John B., "Consumption Taxation in a General Equilibrium Model: How Reliable Are Simulation Results?" in B.D. Bernheim and J.B. Shoven (ed.), *National Saving and Economic Performance* (Chicago: University of Chicago Press, 1991).

Coller, Maribeth; Harrison, Glenn W., and Rutström, E. Elisabet, "Are Discount Rates Constant? Reconciling Theory and Observation," *Working Paper 3-31*, Department of Economics, College of Business Administration, University of Central Florida, 2003.

Coller, Maribeth, and Williams, Melonie B., "Eliciting Individual Discount Rates," *Experimental Economics*, 2, 1999, 107-127.

Cox, James C., and Sadiraj, Vjollca, "Implications of Small- and Large-Stakes Risk Aversion for Decision Theory," *Unpublished Manuscript*, Department of Economics, University of Arizona, May 2004.

Donkers, Bas, and van Soest, Arthur, "Subjective Measures of Household Preferences and Financial Decisions," *Journal of Economic Psychology*, 20(6), 1999, 613-642.

Eckel, Catherine; Johnson, Cathleen, and Montmarquette, Claude, "Savings Decisions of the Working Poor: Short- and Long-term Horizons," in J. Carpenter, G.W. Harrison and J.A. List (eds.), *Field Experiments in Economics* (Greenwich, CT: JAI Press, Research in Experimental Economics, Volume 10, 2005).

Frederick, Shane; Loewenstein, George; and O'Donoghue, Ted, "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature*, 40, June 2002, 351-401.

Gollier, Christian, *The Economics of Risk and Time* (Cambridge, MA: MIT Press, 2001).

Harrison, Glenn W., Harstad, Ronald M., and Rutström, E. Elisabet, "Experimental Methods and Elicitation of Values," *Experimental Economics*, 7(2), June 2004, 123-140.

Harrison, Glenn W.; Jensen, Jesper; Lau, Morten Igel, and Rutherford, Thomas F., "Policy Reform Without Tears," in A. Fossati and W. Weigard (eds.), *Policy Evaluation With Computable General Equilibrium Models* (New York: Routledge, 2002).

Harrison, Glenn W.; Johnson, Eric; McInnes, Melayne M., and Rutström, E. Elisabet, "Risk Aversion and Incentive Effects: Comment," *Working Paper 3-19*, Department of Economics, College of Business Administration, University of Central Florida, 2003; forthcoming, *American Economic Review*.

Harrison, Glenn W.; Lau, Morten Igel; Rutström, E. Elisabet, and Sullivan, Melonie B., "Eliciting Risk and Time Preferences Using Field Experiments: Some Methodological Issues," in J. Carpenter, G.W. Harrison and J.A. List (eds.), *Field Experiments in Economics* (Greenwich, CT: JAI Press, Research in Experimental Economics, Volume 10, 2005).

Harrison, Glenn W.; Lau, Morten Igel, and Williams, Melonie B., "Estimating Individual Discount Rates for Denmark: A Field Experiment," *American Economic Review*, 92(5), December 2002, 1606-1617.

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

Harrison, Glenn W.; Rutherford, Thomas F., and Tarr, David G., "Trade Liberalization, Poverty and Efficient Equity," *Journal of Development Economics*, 71, June 2003, 97-128.

Harrison, Glenn W.; Rutherford, Thomas F., Tarr, David G., and Gurgel, Antonio, "Trade Policy and Poverty Reduction in Brazil," *World Bank Economic Review*, 18(3), 2004.

Harrison, Glenn W., and Vinod, H.D., "The Sensitivity Analysis of Applied General Equilibrium Models: Completely Randomized Factorial Sampling Designs," *Review of Economics and Statistics*, 74, May 1992, 357-362.

Hartog, Joop; Ferrer-i-Carbonell, Ada, and Jonker, Nicole, "Linking Measured Risk Aversion to Individual Characteristics," *Kyklos*, 55, 2002, 3-26.

Hey, John D., "Experimental Economics and the Theory of Decision Making Under Uncertainty," *Geneva Papers on Risk and Insurance Theory*, 27(1), June 2002, 5-21.

Holt, Charles A., and Laury, Susan K., "Risk Aversion and Incentive Effects," *American Economic Review*, 92(5), December 2002, 1644-1655.

Kahneman, Daniel; Knetsch, Jack L., and Thaler, Richard H., "Experimental Tests of the Endowment Effect and the Coase Theorem," *Journal of Political Economy*, 98, December 1990, 1325-1348.

Kapteyn, Arie, and Teppa, Federica, "Hypothetical Intertemporal Consumption Choices," *Economic Journal*, 113, March 2003, C140-C151.

Kirby, Kris N., and Maraković, Nino N., "Delay-discounting probabilistic rewards: Rates decrease as amounts increase," *Psychonomic Bulletin & Review*, 1996, 3(1), 100-104.

Kirby, Kris N.; Petry, Nancy M., and Bickel, Warren K., "Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls," *Journal of Experimental Psychology: General*, 1999, 128(1), 78-87.

Levin, Irwin P.; Snyder, Mary A., and Chapman, Daniel P., "The Interaction of Experiential and Situational Factors and Gender in a Simulated Risky Decision-Making Task," *Journal of Psychology*, 122(2), March 1988,173–181.

Lau, Morten Igel, "Assessing Tax Reforms When Human Capital is Endogenous," in G. W. Harrison, S. E. H. Jensen, L. H. Pedersen and T. F. Rutherford (eds.), *Using Dynamic General Equilibrium Models for Policy Analysis* (Amsterdam: North Holland, Contributions to Economic Analysis 248, 2000).

Liang, K-Y., and Zeger, S.L., "Longitudinal Data Analysis Using Generalized Linear Models," *Biometrika*, 73, 1986, 13-22.

Markowitz, Harry, "The Utility of Wealth," *Journal of Political Economy*, 60, April 1952, 151-158.

Mitchell, Robert C., and Carson, Richard T., *Using Surveys to Value Public Goods: The Contingent Valuation Method* (Baltimore: Johns Hopkins Press, 1989).

Palacios-Huerta, Ignacio; Serrano, Roberto, and Volij, Oscar, "Rejecting Small Gambles Under Expected Utility," *Working Paper*, Brown University, August 2002.

Piggott, John, and Whalley, John, *UK Tax Policy and Applied General Equilibrium Analysis* (New York: Cambridge University Press, 1985).

Quizon, Jaime B.; Binswanger, Hans P., and Machina, Mark J., "Attitudes Toward Risk: Further Remarks," *Economic Journal*, 94, March 1984, 144-148.

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

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

Robson, Arthur J., "The Evolution of Attitudes to Risk: Lottery Tickets and Relative Wealth," *Games and Economic Behavior*, 14, 1996, 190-207.

Rogers, W. H., "Regression standard errors in clustered samples," *Stata Technical Bulletin*, 13, 1993, 19-23.

Rubin, P. H., and Paul, C. W., "An Evolutionary Model of Taste for Risk," *Economic Inquiry*, 17, 1979, 585–595.

Rubinstein, Ariel, "Comments on the Risk and Time Preferences in Economics," *Unpublished Manuscript*, Department of Economics, Princeton University, 2002.

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

Samuelson, Paul A., "Probability, Utility, and the Independence Axiom," *Econometrica*, 20, 1952, 670-678.

van Praag, Bernard M.S., and Booij, Adam S., "Risk Aversion and the Subjective Time Discount Rate: A Joint Approach," *Working Paper*, Department of Economics and Econometrics, University of Amsterdam, July 2003.

Watt, Richard, "Defending Expected Utility Theory," *Journal of Economic Perspectives*, 16(2), Spring 2002, 227-228.

Williams, Rick L., "A Note on Robust Variance Estimation for Cluster-Correlated Data," *Biometrics*, 56, June 2000, 645-646.