Scand. J. of Economics 109(2), 341–368, 2007
DOI: 10.1111/j.1467-9442.2007.00496.x

Estimating Risk Attitudes in Denmark: A Field Experiment*

Glenn W. Harrison

University of Central Florida, Orlando, FL 32816, USA
gharrison@research.bus.ucf.edu

Morten I. Lau

Durham University, Durham DH1 3LB, England
m.i.lau@durham.ac.uk

E. Elisabet Rutström

University of Central Florida, Orlando, FL 32816, USA
erutstrom@bus.ucf.edu

Abstract

We estimate individual risk attitudes using controlled experiments in the field in Denmark. 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. 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 such as age and education. However, we do not find any effect of sex on risk attitudes.

Keywords: Risk aversion; field experiment; socio-demographic characteristics

JEL classification: *C*81; *C*93; *D*01; *D*81

I. Introduction

Welfare analysis of major economic policy decisions such as those concerning education, employment, healthcare, 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

---

*Rutström thanks the U.S. National Science Foundation for research support under grants NSF/IIS 9817518 and NSF/POWRE 9973669, and Harrison and Rutström are grateful for support under grants NSF/HSD 0527675 and NSF/SES 0616746. Harrison and Lau thank the Danish Social Science Research Council for research support under project no. 24-02-0124. Steffen Andersen provided superb research assistance and comments throughout. We are grateful to Christian Gollier, John Hey, David Levine, Chris Starmer, Melonie Sullivan, Peter Wakker and three referees for comments. Supporting data, statistical code and experimental instructions are stored in the ExLab Digital Archive at http://exlab.bus.ucf.edu.

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.

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. 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 the type of income changes considered here.

We use choices over lotteries with real monetary rewards to elicit risk attitudes. The lottery choices are based on those used by Holt and Laury (2002), who elicited risk attitudes for university students using 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.[1] 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, 2005) and Harrison, Johnson, McInnes and Rutström (hereafter HJMR, 2005) 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 education affects risk attitudes, and that those who begin or complete vocational training or higher education are significantly more risk averse than those with less. We also find an effect from age, where those in our middle-aged group (between ages 40 and 50) are less risk averse than others. This is an effect that would simply not be

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

observable using subjects typically recruited on college campuses, since these subject pools do not include older people. However, we do not find any effect of sex on risk attitudes. This is noteworthy because many previous studies using similar experimental and survey methods suggest that women are more risk averse than men.

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.* We also do not find evidence that any of the demographic sub-groups considered exhibit a relative risk aversion that changes over the income domain considered here. Thus, *CRRA is an appropriate assumption for all the identifiable sub-groups of the Danish population and the lottery stakes considered here.*

At a methodological level, we show that it is possible to carry out an elicitation of risk attitudes in the field in order to generate a wider demographic variation than one would get in the lab, therefore better reflecting the population of a country. 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.[2]

In Section II we review our experimental design. We propose several extensions of the basic 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 III explains the field experiments conducted, with additional details on procedures provided in Harrison, Lau, Rutström and Sullivan (hereafter HLRS, 2005). Our design allows one to identify possible framing effects, and evaluate relatively flexible functional forms for risk preferences. Section IV examines the results and relates them to those found in the existing literature.

[2] In the Epilogue to a book-length review of the economics of risk and time, Gollier (2001, pp. 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. 425–426) about the long-standing need for empirical evaluations of restrictive functional forms such as CRRA. In an appendix we discuss a number of issues surrounding the interpretation of estimates of risk aversion from experiments and the validity of expected utility theory.

II. Experimental Design

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) and HJMR (2005).[3] 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 2,000 DKK and a 90% chance of receiving 1,600 DKK. The expected value of this lottery, EV^A, is shown in the third-last column as 1,640 DKK, although the EV columns were not presented to subjects. Similarly, lottery B in the first row has chances of payoffs of 3,850 DKK and 100 DKK, for an expected value of 480 DKK. Thus the two lotteries have a relatively large difference in expected values, in this case 1,170 DKK.[4] 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 either the *highest* or the *lowest* row in Table 1 at which the subject "switched" over from lottery A to lottery B. 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, the data could be treated as a panel with each subject making 10 choices, one for each row in the table. Finally, one could study the effects of experimental conditions in terms

[3] 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). Andersen, Harrison, Lau and Rutström (2006) compare behavior using a standard MPL to the extended MPL used here.

[4] At the time of the experiment the exchange rate was approximately 6.55 DKK 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.

Table 1. *Typical payoff table for risk-aversion experiments*

Lottery A				Lottery B							
p(2,000)		p(1,600)		p(3,850)		p(100)		EV^A	EV^B	Difference	Open CRRA interval if subject switches to lottery B
0.1	2,000	0.9	1,600	0.1	3,850	0.9	100	1,640	480	1,170	$-\infty$, -1.71
0.2	2,000	0.8	1,600	0.2	3,850	0.8	100	1,680	850	830	-1.71, -0.95
0.3	2,000	0.7	1,600	0.3	3,850	0.7	100	1,720	1,230	490	-0.95, -0.49
0.4	2,000	0.6	1,600	0.4	3,850	0.6	100	1,760	1,600	160	-0.49, -0.15
0.5	2,000	0.5	1,600	0.5	3,850	0.5	100	1,800	1,980	−170	-0.14, 0.14
0.6	2,000	0.4	1,600	0.6	3,850	0.4	100	1,840	2,350	−510	0.15, 0.41
0.7	2,000	0.3	1,600	0.7	3,850	0.3	100	1,880	2,730	−840	0.41, 0.68
0.8	2,000	0.2	1,600	0.8	3,850	0.2	100	1,920	3,100	−1,180	0.68, 0.97
0.9	2,000	0.1	1,600	0.9	3,850	0.1	100	1,960	3,480	−1,520	0.97, 1.37
1	2,000	0	1,600	1	3,850	0	100	2,000	3,850	−1,850	1.37, ∞

Notes: All currency units are Danish kroner (DKK). At the time of the experiment 1 USD = 6.55 DKK. The last four columns in this table, showing the expected values of the lotteries and the implied CRRA intervals, were not shown to subjects.

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.[5] The dependent variable in the interval regression model is the CRRA interval that subjects implicitly choose when they switch from lottery 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 five 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 seven safe choices would have revealed a CRRA interval between 0.68 and 0.97, and so on.

The CRRA characterization of risk attitudes is popular in theoretical and applied work, no doubt due to its 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 for a subset of that domain. To allow for the possibility that the relative risk aversion is not constant one may also estimate a flexible functional form using maximum likelihood, such as the expo-power function proposed by Saha (1993). The expo-power 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 CARA (as r tends to 0), but is not defined for α equals 0.

Extensions

We expanded 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

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

that there were some responses 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: 2,000, 1,600; B1: 3,850, 100), (A2: 2,250, 1,500; B2: 4,000, 500), (A3: 2,000, 1,750; B3: 4,000, 150), and (A4: 2,500, 1,000; B4: 4,500, 50). These prizes range from approximately \$7.65 to \$687.[6]

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.[7] 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.

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 options and chooses one for each row. Subjects are generally allowed to switch back and forth as they like. 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. Subjects were also given an explicit option of expressing indifference between the lottery options on each row. One possible explanation for the observation that subjects switch back and forth between choices in MPL is that they are indifferent. If so, explicitly including an indifference option may be a cleaner way to capture this behavior. The second extension of the

[6] Holt and Laury (2002) and HJMR (2005) had subjects make choices in more than one lottery task, where the prizes were simply scaled up by a factor. It is possible that there are behavioral differences between these different ways of varying prizes, but that is not our focus here. These lotteries were simply designed to give us a richer data set from which to statistically identify risk attitudes.

[7] 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, −1.10, −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.

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.[8]

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. Subjects complete all their choices on all relevant tables before any of the random selections are carried out.

As the subject iterates in the iMPL the choices should appear more and more alike, by design. Hence one would expect that greater cognitive effort would be needed to discriminate between the two lottery pairs. At some point we expect the subject to express indifference, which we account for in our analysis by only including observations on the intervals over which the subject made strict choices in the form of one or the other lottery.

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. 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.[9] 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

[8] 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.
[9] Kirby and Maraković (1996), Kirby, Petry and Bickel (1999) and Eckel, Johnson and Montmarquette (2005).

identify statistically. 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 use all available information to identify a good valuation for a commodity whenever there is any kind of uncertainty involved.[10] 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 devised a test for framing effects by varying the cardinal scale of the MPL used in the risk-aversion task. Two asymmetric frames were developed in addition to the standard, symmetric one: 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 six decision rows in Level 1 of the iMPL, as opposed to the 10 rows in the symmetric frame.[11] As suggested by the treatment names, *skewLO* (*skewHI*) is intended to skew responses to be lower (higher) probabilities if subjects pick in the middle.

III. Procedures in Denmark

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, essentially following those employed in Harrison *et al.* (2002). Full details are provided in HLRS (2005):

(i) 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 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.

(ii) Second, we discarded 17 municipalities (including one county) from the population, due to their extraordinarily remote locations. The population represented in these locations amounts to less than 2% of

[10] See Harrison, Harstad and Rutström (2004).

[11] 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 one 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.

the Danish population, or 493 individuals in our sample of 25,000 from the Civil Registry.

(iii) Third, we assigned each county either one session or two 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.

(iv) Fourth, we divided six 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.

(v) 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.

(vi) 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, corresponding to a response rate of 40%.

Attendance at the experimental sessions was extraordinarily high, including four subjects who did not respond to the letter of invitation but showed up unexpectedly and participated in the experiment. Four subjects turned up for their session, but were not able to participate in the experiments.[12] These experiments were conducted in June of 2003, and a total of 253 subjects participated.[13] 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.

[12] 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 he arrived after the instructions had begun and we had already included one unexpected "walk-in" to fill his position.

[13] Certain events might have plausibly triggered some of the no-shows: for example, three 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.

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.[14] 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 3–11 balls, depending on the number of decision rows in the iMPL used in 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 HLRS (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, sex, 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

[14] 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, Ferrer-i-Carbonell and Jonker (2002) and van Praag and Booij (2003). Dohmen, Falk, Huffman, Sunde, Schupp and Wagner (2005) combine hypothetical surveys with experiments that involve real monetary rewards.

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 *et al.* (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 case 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”.

IV. Results

We report our 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?

Risk Aversion

Figure 1 shows the kernel density of observed 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.[15] 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, 2005) and HJMR (2005).[16] 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.[17]

[15] 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.

[16] Our payoffs are roughly equal to 150 times the base payoffs in Holt and Laury (2002), which is considerably larger than the largest one employed by them of 90 times the base. For this payoff factor they find a CRRA of 0.64. Nevertheless, since we only paid each subject with a 10% probability, this could be argued to lower our payoff scale to one of a factor of 15, which is closer to the factor of 20 employed by them. For a factor of 20 they find a CRRA coefficient of 0.51 and for a factor of 50 one of 0.70. HJMR (2005) find a coefficient of 0.57 for a factor of 10. We also conducted a small experiment to directly examine the hypothesis that paying subjects with a 1-in-10 probability generates the same responses as paying them for certain. The control condition was a Holt and Laury (2002) design with lottery payoffs that were \$16, \$20, \$1 and \$38.50, which is 10 times their baseline payoffs. The treatment condition differed only in that we rolled a 10-sided die for each of them individually to determine if they were to be paid. We recruited 77 subjects from the University of Central Florida student population: 51 in the control treatment and 26 in the 1-in-10 treatment. Using interval regression methods to estimate the effect of the treatment on elicited risk attitudes, we find that it is associated with a CRRA that is 0.11 lower than the control, but with a standard error of 0.13. Hence the p-value on the hypothesis test that there is no effect on risk attitudes from paying subjects with a 10% chance is 0.39, and we cannot reject the null hypothesis.

[17] Andersen, Harrison, Lau and Rutström (2005) consider the effect of assuming that the subjects made their choices in accordance with prospect theory, instead of the expected utility theory specification assumed here. Using joint estimation of risk and time preferences, they find that the estimate of the r parameter drops from 0.64 under Expected Utility Theory to 0.55 under Separable Prospect Theory and 0.44 under Cumulative Prospect Theory. These

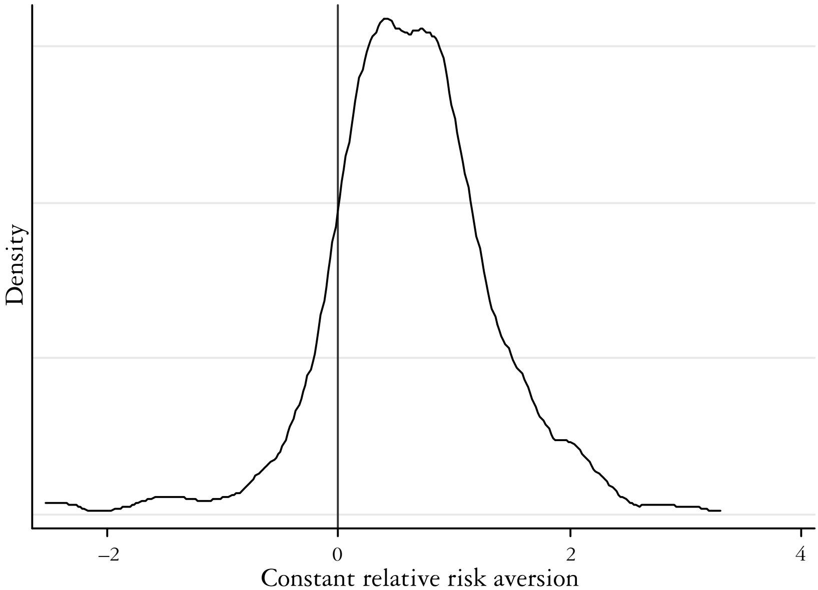

Fig. 1. Distribution of CRRA in Denmark with symmetric menu; midpoint of raw responses from iMPL

The unconditional data indicate that there is an effect on elicited risk aversion from the 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 by the hypothesis that subjects are biased towards the middle of the table.[18] 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

---

parameter values are significantly different from zero and do not change our conclusion that the average Dane is risk averse.

[18] If we use the lower endpoint of the elicited CRRA interval, the population weighted mean values are 0.65 for the symmetric treatment, 0.48 for the *SkewLO* treatment and 0.88 for the *SkewHI* treatment. The raw midpoint of the CRRA interval is not defined if there is no lower or upper endpoint of the elicited CRRA interval, and the number of observations is therefore different when we use interval endpoints instead of midpoints. Thus, even though using lower endpoints should generate lower CRRA estimates, the one for *skewLO* is actually higher here due to the loss of observations. Nevertheless, the important thing to notice is that our main conclusion that Danes are generally risk averse is unaffected.

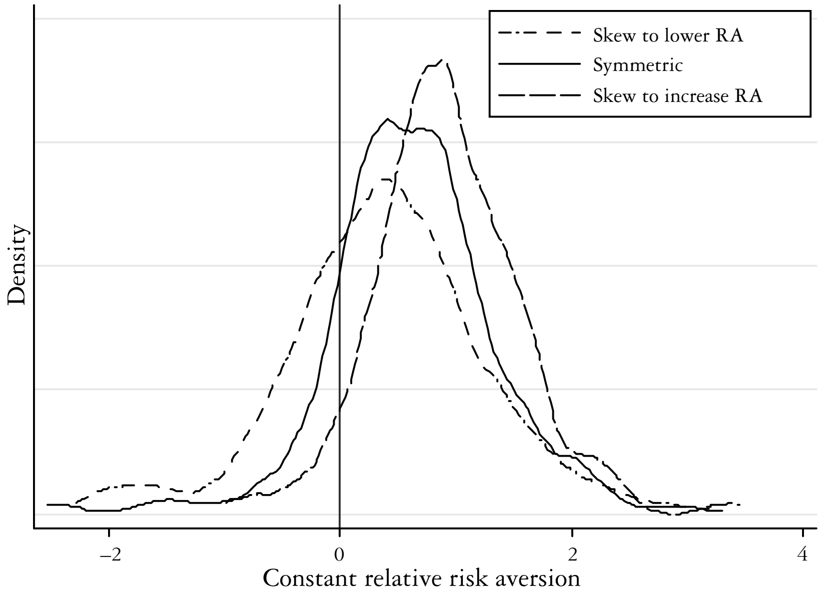

Fig. 2. Effect of initial menu on estimated CRRA in Denmark; midpoint of raw responses from iMPL

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). A conditional analysis also allows us to answer our second question: do risk attitudes vary with observable demographic characteristics?

Constancy of Relative Risk Aversion and Heterogeneity

In order to assess the importance of demographics on risk attitudes, we apply 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. Table 2 shows that the raw sample means of these characteristics are similar to the estimated population means, such that our sample appears representative of the population. The largest difference is for the Aged over 50 group, but it is only five percentage points. To be safe, we also correct for sample selection bias in our analysis of demographic effects.

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
County_15	Koebenhavns Amt	0.10	0.11
County_20	Frederiksborg Amt	0.09	0.07
County_25	Roskilde Amt	0.05	0.04
County_30	Vestsjaellands Amt	0.09	0.10
County_42	Fyns Amt	0.09	0.09
County_50	Soenderjyllands Amt	0.05	0.05
County_55	Ribe Amt	0.06	0.04
County_60	Vejle Amt	0.06	0.07
County_65	Ringkoebing Amt	0.11	0.09
County_70	Aarhus Amt	0.09	0.12
County_80	Nordjyllands Amt	0.10	0.09
wave2	Second wave of invitations	0.03	0.03
wave3	Third wave of invitations	0.02	0.02

Notes: 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 programs 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. The Danish names for each county are provided, where "Amt" means county. Counties 30 and 35 are aggregated, due to smaller samples, as are counties 65 and 76. These are neighboring counties in both cases.

We first consider the hypothesis that relative risk aversion is constant over the income domain considered in our experiments. We then examine the marginal effects of demographics and treatments. Finally, we evaluate the total effects of several of the characteristics, by estimating marginal effects without controls for other characteristics. We calculate total effects since many demographic characteristics co-vary in the population and therefore also in our sample. 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. By not controlling for these other characteristics of men, we can estimate the difference in risk attitudes between men and women that jointly reflects all of these differences.

Constancy of Relative Risk Aversion. Maximum likelihood estimates of the expo-power 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.[19] 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. Responses are coded as the binary choice of each subject on each row of the MPL, and we therefore have a panel. We allow for the responses of the same subject to be correlated, due to unobserved individual effects, and account for the complex 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.

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 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, as in Liang and Zeger (1986), and generalize

[19] 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 allows 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.

the "robust standard errors" approach popular in econometrics, as in Rogers (1993).

The estimates indicate that there is no statistically significant deviation in α from zero for any of the individual characteristics controlled for, or for the constant. We therefore conclude *that there is no evidence to reject CRRA as an appropriate general characterization for this sample and this income domain*. Moreover, we do not identify any sub-group of the Danish population for which the CRRA assumption can be rejected. Based on these findings we conclude that CRRA is an acceptable restriction on the functional form of the utility function. Since the expo-power function is not defined for constant relative risk aversion, we estimate the CRRA function directly to obtain reliable estimates of treatment and demographic effects.

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. We apply regression models that condition on observable characteristics and allow for selection biases using techniques standard in econometrics.[20] Responses are coded as the interval defined by the CRRA values for which the subject switched from the safe to the risky lottery in the final iteration. The size of the intervals differs across subjects since the number of iterations may have differed across subjects, and some subjects also expressed indifference.

Table 3 displays the results from maximum likelihood estimation of a sample selection model of elicited risk attitudes, as well as a comparable model that does not allow for sample selection. Both sets of estimates allow for the complex survey design, sample weights that reflect the adult population of Denmark and the possibility that observations on choices made by the same subject are not independent. We find evidence of significant sample selection into the experiments. The ancillary parameter ρ measures the estimated correlation between the residuals of the sample selection equation and the main CRRA equation. It equals 0.46, has a standard error of only 0.18, and has a 95% confidence interval with values of $+0.04$ and $+0.74$. If this correlation had been zero there would have been no evidence of sample selection bias on the main estimates of CRRA. The coefficients in the sample selection are jointly significant, as are many of the individual coefficients.

We find three types of characteristics that appear to be correlated with variations in the risk attitudes across subjects. Table 3 shows that the middle-aged group (between ages 40 and 50) are less risk averse than

[20] See Vella (1998) for a review of the range of sample selection correction techniques available. We employ a two-step maximum likelihood estimation of the Heckman (1976, 1979) selection model, with corrections to standard errors for the complex sample survey design.

Table 3. *Statistical model of marginal effects on risk-aversion responses*

		Sample selection correction			No correction		
Variable	Variable description	Estimate	Standard error	*p*-Value	Estimate	Standard error	*p*-Value
A. CRRA equation							
Constant		−0.36	0.29	0.22	0.30	0.28	0.29
skewLO	Frame to skew RA down	−0.16	0.10	0.10	−0.05	0.13	0.71
skewHI	Frame to skew RA up	0.29	0.08	0.00	0.24	0.09	0.01
Task2	Second RA task	0.30	0.06	0.00	0.29	0.06	0.00
Task3	Third RA task	0.23	0.05	0.00	0.20	0.05	0.00
Task4	Fourth RA task	0.18	0.04	0.00	0.20	0.05	0.00
experimenter	Experimenter Steffen Andersen	−0.10	0.08	0.19	−0.04	0.09	0.63
female	Female	−0.01	0.07	0.93	−0.03	0.09	0.78
young	Aged less than 30	0.18	0.17	0.28	0.09	0.20	0.63
middle	Aged between 40 and 50	−0.26	0.12	0.03	−0.29	0.14	0.03
old	Aged over 50	−0.20	0.13	0.13	−0.16	0.17	0.35
single	Lives alone	0.13	0.12	0.28	0.07	0.15	0.66
kids	Has children	−0.02	0.11	0.88	0.01	0.14	0.93
nhhd	Number in household	0.02	0.05	0.68	−0.01	0.07	0.91
owner	Own home or apartment	0.19	0.09	0.04	0.05	0.12	0.67
retired	Retired	0.04	0.11	0.72	−0.10	0.13	0.43
student	Student	0.29	0.14	0.04	0.31	0.17	0.07
skilled	Some post-secondary education	0.27	0.09	0.00	0.29	0.10	0.00
longedu	Substantial higher education	0.35	0.10	0.00	0.38	0.12	0.00
IncLow	Lower-level income	−0.01	0.10	0.96	−0.02	0.13	0.87
IncHigh	Higher-level income	0.00	0.09	0.98	−0.03	0.12	0.79
copen	Lives in Copenhagen area	0.18	0.11	0.10	0.13	0.12	0.28
city	Lives in larger city of 20,000 or more	0.08	0.09	0.35	0.17	0.11	0.12

Continued

Table 3. *(Continued)*

Variable	Variable description	Sample selection correction			No correction		
		Estimate	Standard error	*p*-Value	Estimate	Standard error	*p*-Value
B. Sample selection equation							
Constant		0.75	0.10	0.00			
female	Female	−0.14	0.09	0.14			
young	Aged less than 30	0.13	0.14	0.34			
middle	Aged between 40 and 50	0.22	0.13	0.09			
old	Aged over 50	0.01	0.12	0.96			
County_15	Koebenhavns Amt	−0.24	0.08	0.00			
County_20	Frederiksborg Amt	−0.35	0.09	0.00			
County_25	Roskilde Amt	−0.41	0.11	0.00			
County_30	Vestsjaellands Amt	−0.58	0.09	0.00			
County_42	Fyns Amt	−0.30	0.07	0.00			
County_50	Soenderjyllands Amt	−0.42	0.11	0.00			
County_55	Ribe Amt	−0.52	0.13	0.00			
County_60	Vejle Amt	0.03	0.09	0.71			
County_65	Ringkoebing Amt	−0.05	0.09	0.57			
County_70	Aarhus Amt	−0.32	0.08	0.00			
County_80	Nordjyllands Amt	−0.40	0.09	0.00			
wave2	Second wave of invitations	−0.39	0.23	0.09			
wave3	Third wave of invitations	−0.07	0.39	0.86			
ρ	Error correlation	0.46	0.18				

Notes: Survey interval regression estimates of CRRA corrected for sample selection bias. $N = 925$ choices by 245 subjects stratified across 12 counties.

both those who are older and those who are younger. The significance level compared to the reference group, ages 30 to 39, is 3.2%. We also find an effect from being a student or completing vocational training or higher education that raises risk aversion. Students and individuals with some post-secondary or substantial higher education are associated with higher risk attitudes than those with less education.

There is no significant effect from sex on risk attitudes. The absence of an effect of sex is noteworthy, since it has been intensively studied using related experimental and survey methods, and has even been the focus of theorizing about the role of evolution in forming preferences.[21]

The results in Table 3 also suggest that there are some significant effects on risk aversion from our experimental treatments, as well as from each of the four tasks. In particular, Task 2 is associated with higher CRRA responses, with a significant coefficient value of 0.30. The tasks presented to each subject differ in the prizes used, although the prizes are not simply varied in a monotonic manner. Since we reject the possibility of non-constancy in the relative risk aversion in our expo-power estimation, we conclude that these task effects do not simply reflect income effects. In a series of complementary lab experiments we find that these task effects partly reflect an order effect, such that later tasks are more likely to lead to higher RRA than earlier ones. The lab experiments also found a significant difference in choices due to the particular prize structure employed in the second task, independent of the order.[22] It is important to control for task when we investigate our treatment effects so as not to confound those tests.

The variables *skewLO* and *skewHI* control for the frame used. The *skew-HI* treatment is statistically significant, but there is no significant effect

[21] Levin, Snyder and Chapman (1988) and Powell and Ansic (1997) illustrate the experimental studies undertaken in a settings in which the task was not abstract but there were no real earnings by subjects. Harbaugh, Krause and Vesterlund (2002) and Holt and Laury (2002) conduct abstract experiments with real rewards, and find no significant sex effects on elicited risk aversion when stakes are non-trivial. Schubert, Brown, Gysler and Brachinger (1999) conduct abstract and non-abstract experiments with real rewards, and conclude that women do appear to be more risk averse than men in abstract tasks in the gain frame, but that this effect disappears with context. Unfortunately, they employed the Becker–DeGroot–Marschak procedure for eliciting certainty-equivalents, which is known to have poor incentive properties for experimental subjects; see Harrison (1992). Jianakoplos and Bernasek (1998) examine data from the *U.S. Survey of Consumer Finances*, and conclude that single women are more risk averse in their financial choices than single men. Rubin and Paul (1979) and Robson (1996) offer evolutionary models of possible sex differences in risk aversion. Even if there is no evidence for an effect of sex on risk aversion, it is possible that observers may *predict* differences in risk attitudes based on sex; see Eckel and Grossman (2002).

[22] The lab results are reported in Andersen *et al.* (2006). HJMR (2005) also report the presence of order effects in these kind of lottery choice tasks. To separate order from stake effects for the Danish data we varied the order of the four tasks in our complementary laboratory experiments.

from the *skewLO* treatment.[23] The *skewLO* treatment yields an average CRRA of 0.43, and the *skewHI* treatment yields an average CRRA of 0.91, each in a direction expected *a priori*. Both estimates are consistent with the conclusion that subjects in the field experiments are risk averse.

Total Effects. Now consider the total, rather than marginal, effects of key demographic variables. One example mentioned earlier is how the total effects may differ from the marginal effects: 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 to 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. To consider the total effects, we simply repeat the statistical analysis shown in Table 3 (including the sample selection bias correction), 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. The coefficients from each regression are displayed in Table 4.

We find that sex still has no effect on risk aversion. The only significant total effects come from the same sources as the significant marginal effects: age and education. Younger subjects (up to age 39) are generally more risk averse than the older ones, and the lower risk aversion amongst the older group (above 50) is stronger here as a total effect than it was as a marginal effect. Students are even more risk averse when we consider the total effect, most likely because of their younger age. The strong positive marginal effect on risk aversion from the skilled and higher education categories remain, indicating greater risk aversion with the completion of some vocational training or higher education. Finally, the positive marginal effect from living in the Copenhagen area is stronger when we consider the total effect, probably because a larger proportion of young and single people live in the city than otherwise.

These results all clearly show the importance of accounting for heterogeneity in risk attitudes. One important implication is that one might have to evaluate welfare policies at a disaggregated household level in order to correctly identify differences in risk attitudes and hence the correct

[23] HLRS (2005) show that there is a significant effect on elicited risk attitudes from *both* skewness treatments on the *initial* stage of the iMPL, but that the iterations of the iMPL make that effect disappear for the *skewLO* treatment. Our analysis focuses on the final stage of the iMPL procedure. Thus, one should be concerned about the possible effects of such framing on eliciting risk attitudes if using the original MPL procedure.

Table 4. *Statistical model of total effects on risk-aversion responses*

Variable	Description	Estimate	Standard error	p-Value	Lower 95% confidence interval	Upper 95% confidence interval
female	Female	0.05	0.08	0.47	−0.09	0.20
young	Aged less than 30	0.09	0.14	0.49	−0.18	0.36
middle	Aged between 40 and 50	−0.30	0.12	0.01	−0.54	−0.07
old	Aged over 50	−0.28	0.11	0.01	−0.49	−0.07
single	Lives alone	0.13	0.10	0.18	−0.06	0.32
kids	Has children	0.04	0.08	0.65	−0.12	0.20
nhhd	Number in household	0.00	0.03	0.92	−0.07	0.06
owner	Own home or apartment	0.00	0.08	0.98	−0.17	0.16
retired	Retired	−0.17	0.10	0.08	−0.36	0.02
student	Student	0.34	0.13	0.01	0.09	0.59
skilled	Some post-secondary education	0.21	0.10	0.04	0.01	0.40
longedu	Substantial higher education	0.30	0.10	0.00	0.11	0.50
IncLow	Lower-level income	−0.05	0.09	0.55	−0.23	0.12
IncHigh	Higher-level income	−0.01	0.10	0.93	−0.20	0.18
copen	Lives in Copenhagen area	0.22	0.11	0.04	0.01	0.44
city	Lives in larger city of 20,000 or more	0.15	0.09	0.10	−0.03	0.33

Notes: Survey interval regression estimates of CRRA corrected for sample selection bias. Each coefficient is estimated in a separate model not including other characteristics variables. $N = 842$ choices by 241 subjects stratified across 12 counties. Variables and coefficients that are framed were estimated simultaneously. All others were estimated separately. All estimates have been corrected for sample selection bias.

certainty-equivalent of uncertain policies. For example, if one just compared men and women there would be no difference in risk attitudes: the total effect of sex is effectively zero. But if one compared a young, highly educated Danish man who lived alone, you would find a significantly more risk-averse individual than the average Dane. This conclusion is of more significance for policy analysis than might be initially appreciated. It is common to find household disaggregations 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 several of these major characteristics are varied together, 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.

V. Conclusions

We elicited 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 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, and those who begin or complete vocational training or a higher education are significantly more risk averse than those with less education. We do not, however, find any effect of sex on risk attitudes. Our findings have important implications for the characterization of risk attitudes in policy applications, theoretical modeling and experimental economics.

Appendix. 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-stake 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 $10/11 \times 10/11 \approx 5/6$ 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.[24]

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, original emphasis):

> 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 its 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. Rubinstein (2002) and Cox and Sadiraj (2006) make these points quite clearly. Cox and Sadiraj (2006) 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.

[24] 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 and Serrano (2006) 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 turn down more generous gambles at higher stakes.

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 its own controversies; see Frederick, Loewenstein and O'Donoghue (2002) and Coller, Harrison and Rutström (2003).

Fudenberg and Levine (2006) propose a dual-self model that seeks to reconcile these two perspectives with stylized experimental evidence on risk and time preferences. Their model posits that the decision-maker has two selves.[25] In effect, a latent EUTw-consistent self constrains choices actually observed by the EUTi-motivated self. Under some circumstances observed choices will be EUTi-consistent, and under other circumstances observed choices will be EUTw-consistent. Andersen *et al.* (2005) examine the implications of this framework for experimental data.

References

Andersen, S., Harrison, G. W., Lau, M. I. and Rutström, E. E. (2005), Eliciting Risk and Time Preferences, Working Paper no. 05-24, Department of Economics, College of Business Administration, University of Central Florida.

Andersen, S., Harrison, G. W., Lau, M. I. and Rutström, E. E. (2006), Elicitation using Multiple Price Lists, *Experimental Economics 9*, 383–405.

Benhabib, J. and Bisin, A. (2005), Modeling Internal Commitment Mechanisms and Self-control: A Neuroeconomics Approach to Consumption–Saving Decisions, *Games and Economic Behavior 52*, 460–492.

Coller, M. and Williams, M. B. (1999), Eliciting Individual Discount Rates, *Experimental Economics 2*, 107–127.

Coller, M., Harrison, G. W. and Rutström, E. E. (2003), Are Discount Rates Constant? Reconciling Theory and Observation, Working Paper no. 3-31, Department of Economics, College of Business Administration, University of Central Florida.

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

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J. and Wagner, G. G. (2005), Individual Risk Attitudes: New Evidence from a Large Representative, Experimentally-validated Survey, Discussion Paper no. 1730, Institute for the Study of Labor, Bonn.

Donkers, B. and van Soest, A. (1999), Subjective Measures of Household Preferences and Financial Decisions, *Journal of Economic Psychology 20*, 613–642.

Eckel, C. and Grossman, P. J. (2002), Sex Differences and Statistical Stereotyping in Attitudes towards Financial Risk, *Evolution and Human Behavior 23*, 281–295.

Eckel, C., Johnson, C. and Montmarquette, C. (2005), 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*, JAI Press, Greenwich, NY.

[25] A similar model is also proposed by Benhabib and Bisin (2005). The concept of "dual selves" has a long lineage in behavioral economics.

Frederick, S., Loewenstein, G. and O'Donoghue, T. (2002), Time Discounting and Time Preference: A Critical Review, *Journal of Economic Literature 40*, 351–401.
Fudenberg, D. and Levine, D. K. (2006), A Dual Self Model of Impulse Control, *American Economic Review 96*, 1449–1476.
Gollier, C. (2001), *The Economics of Risk and Time*, MIT Press, Cambridge, MA.
Harbaugh, W. T., Krause, K. and Vesterlund, L. (2002), Risk Attitudes of Children and Adults: Choices over Small and Large Probability Gains and Losses, *Experimental Economics 5*, 53–84.
Harrison, G. W. (1992), Theory and Misbehavior of First-price Auctions: Reply, *American Economic Review 82*, 1426–1443.
Harrison, G. W. and List, J. A. (2004), Field Experiments, *Journal of Economic Literature 42*, 1013–1059.
Harrison, G. W., Harstad, R. M. and Rutström, E. E. (2004), Experimental Methods and Elicitation of Values, *Experimental Economics 7*, 123–140.
Harrison, G. W., Lau, M. I. and Williams, M. B. (2002), Estimating Individual Discount Rates for Denmark: A Field Experiment, *American Economic Review 92*, 1606–1617.
Harrison, G. W., Johnson, E., McInnes, M. M. and Rutström, E. E. [HJMR] (2005), Risk Aversion and Incentive Effects: Comment, *American Economic Review 95*, 897–901.
Harrison, G. W., Lau, M. I., Rutström, E. E. and Sullivan, M. B. [HLRS] (2005), 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*, JAI Press Greenwich, NY.
Hartog, J., Ferrer-i-Carbonell, A. and Jonker, N. (2002), Linking Measured Risk Aversion to Individual Characteristics, *Kyklos 55*, 3–26.
Heckman, J. J. (1976), The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models, *Annals of Economic and Social Measurement 5*, 475–492.
Heckman, J. J. (1979), Sample Selection Bias as a Specification Error, *Econometrica 47*, 153–161.
Hey, J. D. (2002), Experimental Economics and the Theory of Decision Making under Uncertainty, *Geneva Papers on Risk and Insurance Theory 27*, 5–21.
Holt, C. A. and Laury, S. K. (2002), Risk Aversion and Incentive Effects, *American Economic Review 92*, 1644–1655.
Holt, C. A. and Laury, S. K. (2005), Risk Aversion and Incentive Effects: New Data Without Order Effects, *American Economic Review 95*, 902–912.
Jianakoplos, N. A. and Bernasek, A. (1998), Are Women More Risk Averse?, *Economic Inquiry 36*, 620–630.
Kahneman, D., Knetsch, J. L. and Thaler, R. H. (1990), Experimental Tests of the Endowment Effect and the Coase Theorem, *Journal of Political Economy 98*, 1325–1348.
Kapteyn, A. and Teppa, F. (2003), Hypothetical Intertemporal Consumption Choices, *Economic Journal 113*, C140–C151.
Kirby, K. N. and Maraković, N. N. (1996), Delay-discounting Probabilistic Rewards: Rates Decrease as Amounts Increase, *Psychonomic Bulletin and Review 3*, 100–104.
Kirby, K. N., Petry, N. M. and Bickel, W. K. (1999), Heroin Addicts have Higher Discount Rates for Delayed Rewards than Non-drug-using Controls, *Journal of Experimental Psychology: General 128*, 78–87.
Levin, I. P., Snyder, M. A. and Chapman, D. P. (1988), The Interaction of Experimental and Situational Factors and Gender in a Simulated Risky Decision-making Task, *Journal of Psychology 122*, 173–181.
Liang, K.-Y. and Zeger, S. L. (1986), Longitudinal Data Analysis using Generalized Linear Models, *Biometrika 73*, 13–22.

Markowitz, H. (1952), The Utility of Wealth, *Journal of Political Economy 60*, 151–158.
Mitchell, R. C. and Carson, R. T. (1989), *Using Surveys to Value Public Goods: The Contingent Valuation Method*, Johns Hopkins University Press, Baltimore, MD.
Palacios-Huerta, I. and Serrano, R. (2006), Rejecting Small Gambles under Expected Utility, *Economics Letters 91*, 250–259.
Powell, M. and Ansic, D. (1997), Gender Differences in Risk Behavior in Financial Decision-making: An Experimental Analysis, *Journal of Economic Psychology 18*, 605–628.
Quizon, J. B., Binswanger, H. P. and Machina, M. J. (1984), Attitudes toward Risk: Further Remarks, *Economic Journal 94*, 144–148.
Rabin, M. (2000), Risk Aversion and Expected Utility Theory: A Calibration Theorem, *Econometrica 68*, 1281–1292.
Rabin, M. and Thaler, R. (2001), Anomalies: Risk Aversion, *Journal of Economic Perspectives 15*, 219–232.
Robson, A. J. (1996), The Evolution of Attitudes to Risk: Lottery Tickets and Relative Wealth, *Games and Economic Behavior 14*, 190–207.
Rogers, W. H. (1993), Regression standard errors in clustered samples, *Stata Technical Bulletin 13*, 19–23.
Rubin, P. H. and Paul, C. W. (1979), An Evolutionary Model of Taste for Risk, *Economic Inquiry 17*, 65–67.
Rubinstein, A. (2002), Comments on the Risk and Time Preferences in Economics, unpublished manuscript, Department of Economics, Princeton University.
Saha, A. (1993), Expo-power Utility: A Flexible Form for Absolute and Relative Risk Aversion, *American Journal of Agricultural Economics 75*, 905–913.
Samuelson, P. A. (1952), Probability, Utility and the Independence Axiom, *Econometrica 20*, 670–678.
Schubert, R., Brown, M., Gysler, M. and Brachinger, H. W. (1999), Financial Decision Making: Are Women Really More Risk Averse?, *American Economic Review 89*, 381–385.
van Praag, B. M. S. and Booij, A. S. (2003), Risk Aversion and the Subjective Time Discount Rate: A Joint Approach, working paper, Department of Economics and Econometrics, University of Amsterdam.
Vella, F. (1998), Estimating Models with Sample Selection Bias: A Survey, *Journal of Human Resources 33*, 127–169.
Watt, R. (2002), Defending Expected Utility Theory, *Journal of Economic Perspectives 16*, 227–228.
Williams, R. L. (2000), A Note on Robust Variance Estimation for Cluster-correlated Data, *Biometrics 56*, 645–646.

First version submitted August 2005;
final version received May 2006.