Risk Attitudes, Randomization to Treatment, and Self-Selection Into Experiments

by

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

January 2005

Abstract. Randomization to treatment is fundamental to statistical control in the design of experiments. But randomization implies some uncertainty about treatment condition, and individuals differ in their preferences towards taking on risk. Since human subjects often volunteer for experiments, or are allowed to drop out of the experiment at any time if they want to, it is possible that the sample observed in an experiment might be biased because of the risk of randomization. On the other hand, the widespread use of a guaranteed show-up fee that is non-stochastic may generate sample selection biases of the opposite direction, encouraging more risk averse samples into experiments. We undertake a field experiment to directly test these hypotheses that risk attitudes play a role in sample selection. We follow standard procedures in the social sciences to recruit subjects to an experiment in which we measure their attitudes to risk. We exploit the fact that we know certain characteristics of the population sampled, adults in Denmark, allowing a statistical correction for sample selection bias using standard methods. We also utilize the fact that we have a complex sampling design to provide better estimates of the target population. Our results suggest that randomization bias is not a major empirical problem for field experiments of the kind we conducted if the objective is to identify marginal effects of sample characteristics. However, there is evidence that the use of show-up fees may have generated a sample that was more risk averse than would otherwise have been observed.

* 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 also grateful to Richard Hofler for comments. Supporting data, statistical code and output are stored in the public ExLab Digital Library at http://exlab.bus.ucf.edu in project "Self-Selection Into Experiments."

Randomization to treatment is fundamental to statistical control in the design of experiments. But randomization implies some uncertainty about treatment condition, and individuals differ in their preferences towards taking on risk. Since human subjects often volunteer for experiments, or are allowed to drop out of the experiment at any time if they want to, it is possible that the sample observed in an experiment might be biased because of the risk of randomization.[1] In the worst case, subjects in experiments might be those that are least averse to being exposed to risk. For many experiments of biological response this might not be expected to have any influence on measurement of treatment efficacy, although many lab, field and social experiments measure treatment efficacy in ways that could be directly affected by randomization bias.[2]

We undertake a field experiment to directly test the hypothesis that risk attitudes play a role in sample selection.[3] We follow standard procedures in the social sciences to recruit subjects to an experiment in which we measure their attitudes to risk. We exploit the fact that we know certain characteristics of the population sampled, adults in Denmark in 2003, allowing a correction for sample selection bias using well-known methods from econometrics. We also utilize the fact that we have a complex sampling design to provide better estimates of the target population.

The classic problem of sample selection refers to possible recruitment biases, such that the

[1] There is some evidence that other animal species behave as if they have aversion to risk, or prefer risk, in certain environments (Kagel, Battalio and Green [1995; ch.6]). Our focus here is on humans.

[2] Heckman and Smith [1995; p. 99-101] provide many examples, and coin the expression "randomization bias" for this possible effect. Harrison and List [2004] review the differences between lab, field, social and natural experiments in economics, and all could be potentially affected by randomization bias. Pevnistkaya [2004] and Palfrey and Pevnitskaya [2004] use thought experiments and lab experiments to illustrate how risk attitudes can theoretically affect the mix of bidders in sealed-bid auctions with endogenous entry, and thereby change behavior in the sample of bidders observed in the auction.

[3] Endogenous subject attrition from the experiment can also be informative about subject preferences, since the subject's exit from the experiment indicates that the subject had made a negative evaluation of it. See Diggle and Kenward [1994] and Philipson and Hedges [1998] for discussion of this statistical issue.

observed sample is generated by a process that depends on the nature of the experiment.[4] We have discussed already the possibility that the use of randomization could attract subjects to experiments that are *less* risk averse than the population, if the subjects rationally anticipate the use of randomization. Conversely, the use of guaranteed financial remuneration, common in experiments in economics for participation, could encourage those that are *more* risk averse to participate.[5] It is well known in the field of clinical drug trials that persuading patients to participate in randomized studies is much harder than persuading them to participate in non-randomized studies (e..g., Kramer and Shapiro [1984; p.2742ff.]). The same problem applies to social experiments, as evidenced by the difficulties that can be encountered when recruiting decentralized bureaucracies to administer the random treatment (e.g., Hotz [1992]). For example, Heckman and Robb [1985] note that the refusal rate in one randomized job training program was over 90%.

1. Data

A. The Task

We employ a simple experimental measure for risk aversion introduced by Holt and Laury [2002] and extended by Harrison, Lau, Rutström and Sullivan [2005]. Each subject is presented with a series of choices between two lotteries, which we call A or B. All choices are presented simultaneously to the subject. The first choice involves a 10% chance of receiving $2 and a 90% chance of receiving $1.60. The expected value of this lottery, EV^A, is $1.64. Lottery B in the first

[4] More precisely, the statistical problem is that there may be some unobserved individual effects that cause subjects to be in the observed sample or not, and these effects could be correlated with responses once in the observed sample.

[5] Most experiments offer subjects a fixed show-up fee, currently ranging between $5 and $10 in convenience samples within the United States. Subjects can also expect to earn an uncertain income, and most experimenters mention this possibility without indicating any expected value or bounds, since that quantitative information could generate biases in the task itself as subject try to attain that earnings threshold.

choice has chances of payoffs of $3.85 and $0.10, for an expected value of $0.48. Thus the two lotteries have a relatively large difference in expected values, in this case $1.17. The expected value of both lotteries increases with the other choices in the set presented to subjects, but the expected value of lottery B steadily becomes greater than the expected value of lottery A.

Subjects are typically confronted with ten such choices. 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 choice presented above, and only risk-averse subjects would take lottery A in the last few choices. A subject that is neutral to risk 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 choices and then choose B thereafter.

These data are analyzed using alternative specifications of risk attitudes. A popular parametric form is based on the constant relative risk aversion (CRRA) utility function, defined as $U(y) = (y^{1-r})/(1-r)$, where r is the CRRA coefficient.[6]

We extend the basic task in several ways. First, we "iterate" the choices made by each subject so that we refine the interval at which they switch from A to B. This allows us to focus responses that consist of extremely small CRRA intervals. In fact, we will reduce these responses to their mid-points, and view the subject as providing a point response, since the intervals are so small at the end of this iterative process. Second, each subject is presented with four such lottery pairs, so that CRRA is elicited for the same subject over a wider range of income levels. Thus we have repeat measures for each subject, and use appropriate statistical models for such data. Third, subjects were randomly

[6] With this parameterization, $r = 0$ denotes risk neutral behavior, $r > 0$ denotes risk aversion, and $r < 0$ denotes risk loving. When $r = 1$, $U(m) = \ln(m)$. Harrison, Lau and Rutström [2004] demonstrate that CRRA is an appropriate functional form for the adult Danish population, by nesting it into more general functional forms which they estimate and testing the CRRA restriction directly. There is also evidence that risk attitudes elicited in this manner are temporally stable: see Harrison, Johnson, McInnes and Rutström [2005].

assigned to treatments designed to test if their risk attitudes were affected by the way that the task was "framed," since one might expect that some subjects would simply choose to switch in the middle of a series of lotteries.

B. Sampling Procedures

The sample for the field experiments was designed to be representative of the adult Danish population in 2003. There were six steps in the construction of the sample,[7] 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 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, and hence being very costly to recruit. 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. Hence it is unlikely that this exclusion could quantitatively influence our results on sample selection bias.
- 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

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

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.[8] 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.[9] These experiments were conducted in June of 2003, and a total of 253 subjects participated.[10] 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

---

[8] We control for the recruitment wave to which the subject responded in our statistical analysis of sample selection.

[9] 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.

[10] 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.

the adult population of Denmark.

The initial recruitment letter explained the purpose of the experiment and that it was being conducted by the Ministry of Economic and Business Affairs of the Danish Government. The letter clearly identified that there would be some randomization involved in determining earnings. In translation, the uncertainty was explained as follows:

> **You can win a significant amount**
> To cover travel costs, you will receive 500 kroner at the end of the meeting. Moreover, each participant will have a 10 percent chance of receiving an amount between 50 and 4,500 kroner in the first part of the survey, and this amount will also be paid at the end of the meeting. In the second part of the survey, each participant will have a 10 percent chance of receiving at least 3,000 kroner. A random choice will decide who win the money in both parts of the survey.

Of course, this paragraph also made it clear that there would be a fixed, non-stochastic income of 500 kroner. The earnings referred to as the "first part of the survey" were for the risk aversion task, and the earnings referred to as the "second part of the survey" were for a separate task. Thus we know that subjects should have rationally anticipated the use of randomization in these experiments.

C. 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. 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 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. 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 choice tasks and the randomization procedures in Part II of the experiment.

3. Results

In order to assess the importance of sample selection on risk attitudes, we applied regression models that condition on observable characteristics of the subjects and allow for selection biases using techniques standard in econometrics.[11] Table 1 provides the definitions of the explanatory variables and summary statistics. Table 2 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 allows for the complex survey design. In particular, they adjust estimates for the fact that subjects in one county were selected independently of subjects in

---

[11] See Vella [1998] for a review of the range of techniques available. We employed full information maximum likelihood estimation of the parametric Heckman [1976][1979] selection model, with corrections to standard errors for the complex sample survey design employed. Version 8.1 of *Stata* was employed to undertake the estimation: see StataCorp [2003] for documentation.

other counties, that sample weights for each subject reflect the adult population of Denmark, as well as the possibility of correlation between responses by the same subject.[12]

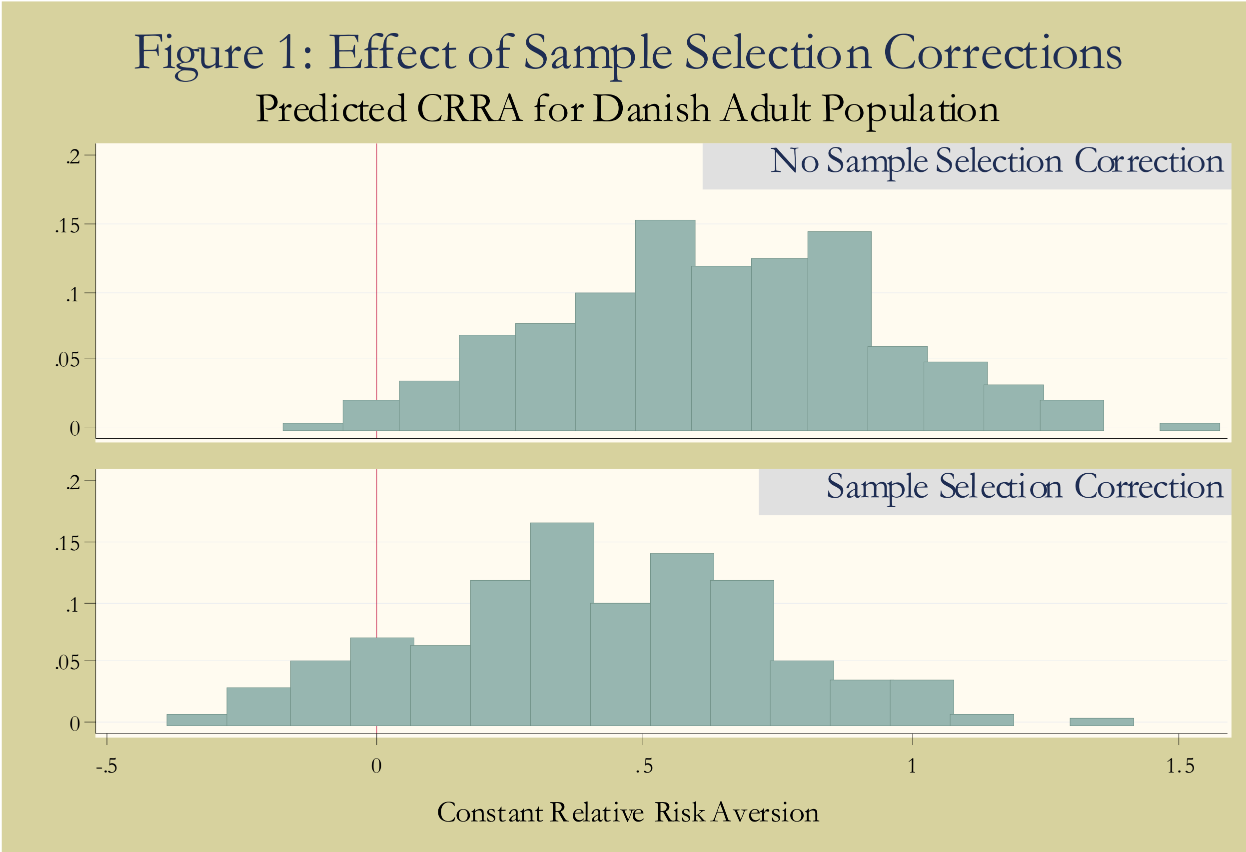
The results indicate that the sample estimates of the main CRRA equation are reliable conditional on the characteristics of the sample observed in the experiment, but that there was evidence of significant sample selection into the experiment. The ancillary parameter ρ measures the estimated correlation between the residuals of the sample selection equation and the main CRRA equation. It equals 0.55, has a standard error of only 0.16, and has a 95% confidence interval with values of +0.15 and +0.80. If this correlation had been zero then 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. On the other hand, the coefficient estimates for the main CRRA equation are virtually identical.[13]

Figure 1 shows histograms of the predicted CRRA from the specification that ignores sample selection bias and then the specification that corrects for it. We find that CRRA is smaller when we correct for sample selection bias, consistent with the hypothesis that use of substantial, guaranteed show-up fees more than offset any bias against attending an experiment that involved randomization. Average CRRA for the specification with no sample selection correction is 0.63, and

---

[12] 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 (see Liang and Zeger [1986]), and generalize the "robust standard errors" approach popular in econometrics (see Rogers [1993]).

[13] Formal tests of pairwise equality for the estimates in the two specifications support this conclusion at any standard significance level. Similarly for a test of the joint hypothesis that all coefficient estimates in the two specifications are the same.



the average for the specification with sample selection corrections is only 0.41. This difference in the central tendency of the predicted distribution of CRRA is significant at standard confidence levels, using parametric or non-parametric tests.

4. Conclusions

Heckman and Smith [1995; p.99] noted that, "Surprisingly, little is known about the empirical importance of randomization bias." Aggregative data about participation rates from job training experiments by Hotz [1992] and clinical trials by Kramer and Shapiro [1984] suggest that it might be significant, but we know of no study that directly evaluates the hypothesis.

Our results suggest that randomization bias is not a major empirical problem for field experiments of the kind we conducted if the objective is to identify marginal effects of sample characteristics. For example, both specifications in Table 2 show that middle-aged Danes have lower risk aversion than others, and the estimated magnitude of this effect is virtually identical. However, there is evidence that the use of certain show-up fees, relatively standard in experimental economics, may have generated a sample that was more risk averse than would otherwise have been observed. Of course, this may be a net effect, since the use of randomization procedures in experiments may have caused an opposite sample selection bias against risk averse subjects choosing to participate.

There are limitations to our analysis. First, one would always like to know "more" about the population being sampled. The Danish environment is a relatively rich one, in which we could identify three major characteristics of the subjects before observing if they agreed to participate. But one could always find that sample selection had no effect on responses by constructing a sufficiently poor statistical model of participation. Since the subjects that do not participate are, by their nature, unobserved from the perspective of the experimenter, this problem is likely to be a general one. Second, our recruitment procedures were typical of those used in economics experiments, in the sense that they used a fixed participation reward which could have offset the effects of randomization. Since this appeared to be a significant factor in our results, it would obviously be interesting in future work to compare the effects of varying the fixed and random component of participation rewards to determine if that influences the risk attitudes of the observed sample.[14] Third, risk attitudes need not be the same for all prizes or outcomes. If our subjects saw the prizes we offered as "small" relative to the implicit prizes from outcomes from other treatments (e.g., life

[14] Rutström [1998] undertakes a design of this kind in the laboratory, varying the show-up fee. She finds considerable differences in the individual characteristics of the subjects that turn up as one varies that fee, but does not conduct a formal test of sample selection biases. Lazear, Malmendier and Weber [2004] design an experiment to test if subjects recruited into a session endogenously sort away from a task that would involve the expression of social preferences, and report significant evidence that they do. Thus their design embeds one sample selection step within an overall experiment, allowing it to be studied intensively. They do not consider sample selection into the overall experiment.

or death in the case of medical interventions), then our results may not generalize since they may have had one risk attitude for "small prizes" and a different one for "large prizes."[15] Fourth, we have not considered the other attrition side of the participation decision for experiments requiring continued enrolment over a period of time or multiple visits. Finally, there is now a rich econometric literature on less parametric specifications of corrections for sample selection, and it would be valuable to see if our results are robust to the use of those specifications.[16]

[15] Such variations in risk attitudes are consistent with economic theory, and are indicated by the controlled experiments of Holt and Laury [2002] and others.

[16] For example, see Das, Newey and Vella [2003] and their references to the literature.

Table 1: 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	0.37
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	0.091
skilled	Some post-secondary education	0.15	0.14
longedu	Substantial higher education	0.36	0.36
IncLow	Lower level income	0.32	0.34
IncHigh	Higher level income	0.36	0.33
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 2: Estimated Relative Risk Aversion

Maximum likelihood estimates, with standard errors corrected for complex survey design

Variable	Variable Description	Sample Selection Correction: Estimate	Sample Selection Correction: Standard Error	Sample Selection Correction: *p*-value	No Correction: Estimate	No Correction: Standard Error	No Correction: *p*-value
		A. CRRA Equation					
Constant		0.03	0.24	0.90	0.25	0.23	0.27
skewLO	Frame to skew RA down	-0.18	0.10	0.07	-0.18	0.10	0.07
skewHI	Frame to skew RA up	0.30	0.08	0.00	0.30	0.08	0.00
experimenter	Experimenter Steffen Andersen	-0.05	0.08	0.48	-0.02	0.08	0.80
female	Female	0.00	0.07	0.99	0.03	0.07	0.67
young	Aged less than 30	0.16	0.17	0.35	0.12	0.17	0.47
middle	Aged between 40 and 50	-0.29	0.12	0.02	-0.33	0.12	0.01
old	Aged over 50	-0.20	0.14	0.14	-0.19	0.14	0.16
single	Lives alone	0.14	0.12	0.23	0.15	0.12	0.21
kids	Has children	-0.02	0.11	0.84	-0.04	0.11	0.73
nhhd	Number in household	0.02	0.05	0.71	0.02	0.05	0.69
owner	Own home or apartment	0.17	0.09	0.06	0.17	0.09	0.07
retired	Retired	0.04	0.11	0.74	0.03	0.11	0.78
student	Student	0.27	0.14	0.05	0.27	0.14	0.06
skilled	Some post-secondary education	0.27	0.09	0.00	0.27	0.09	0.00
longedu	Substantial higher education	0.35	0.10	0.00	0.35	0.10	0.00
IncLow	Lower level income	-0.02	0.10	0.85	-0.02	0.10	0.81
IncHigh	Higher level income	0.01	0.09	0.94	0.01	0.10	0.93
copen	Lives in Copenhagen area	0.13	0.10	0.20	0.08	0.10	0.43
city	Lives in larger city of 20,000+	0.07	0.09	0.44	0.05	0.09	0.54
		B. Sample Selection Equation					
Constant		0.74	0.10	0.00			
female	Female	-0.13	0.09	0.16			
young	Aged less than 30	0.13	0.14	0.33			
middle	Aged between 40 and 50	0.22	0.13	0.09			
old	Aged over 50	0.01	0.12	0.95			
County_15	County 15	-0.25	0.08	0.00			
County_20	County 20	-0.33	0.09	0.00			
County_25	County 25	-0.40	0.11	0.00			
County_30	County 30	-0.57	0.09	0.00			
County_42	County 42	-0.29	0.08	0.00			
County_50	County 50	-0.41	0.11	0.00			
County_55	County 55	-0.51	0.13	0.00			
County_60	County 60	0.02	0.09	0.84			
County_65	County 65	-0.03	0.10	0.73			
County_70	County 70	-0.31	0.09	0.00			
County_80	County 80	-0.42	0.09	0.00			
wave2	Second wave of invitations	-0.41	0.22	0.06			
wave3	Third wave of invitations	-0.08	0.38	0.84			
		C. Ancillary Parameters					
ρ	Error correlation	0.55	0.16				
σ	Standard error of residual in CRRA equation	0.77	0.05				

References

Das, Mitali; Newey, Whitney K., and Vella, Francis, "Nonparametric Estimation of Sample Selection Models," *Review of Economic Studies*, 70, 2003, 33-58.

Diggle, P., and Kenward, M.G., "Informative Drop-out in Longitudinal Data Analysis," *Applied Statistics*, 43(1), 49-93.

Harrison, Glenn W.; Johnson, Eric; McInnes, Melayne M., and Rutström, E. Elisabet, "Temporal Stability of Estimates of Risk Aversion," *Applied Financial Economics Letters*, 2005, 1(1), 31-35.

Harrison, Glenn W.; Lau, Morten I; and Rutström, E. Elisabet, "Estimating Risk Attitudes in Denmark: A Field Experiment," *Working Paper 04-07*, Department of Economics, College of Business Administration, University of Central Florida, 2004.

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, 1009-1055.

Heckman, James J., "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, 1976, 475-492.

Heckman, James J., "Sample Selection Bias as a Specification Error," *Econometrica*, 47, 1979, 153-161.

Heckman, James J., and Robb, Richard, "Alternative Methods for Evaluating the Impact of Interventions," in J. Heckman and B. Singer (eds.), *Longitudinal Analysis of Labor Market Data* (New York: Cambridge University Press, 1985).

Heckman, James J., and Smith, Jeffrey A., "Assessing the Case for Social Experiments," *Journal of Economic Perspectives*, 9(2), Spring 1995, 85-110.

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

Hotz, V. Joseph, "Designing an Evaluation of JTPA," in C. Manski and I. Garfinkel (eds.), *Evaluating Welfare and Training Programs* (Cambridge: Harvard University Press, 1992).

Kagel, John H.; Battalio, Raymond C., and Green, Leonard, *Economic Choice Theory: An Experimental Analysis of Animal Behavior* (New York: Cambridge University Press, 1995).

Kramer, Michael, and Shapiro, Stanley, "Scientific Challenges in the Application of Randomized Trials," *Journal of the American Medical Association*, 252, November 16, 1984, 2739-2745.

Lazear, Edward P.; Malmendier, Ulrike, and Weber, Roberto A., "Sorting in Experiments," *Unpublished Manuscript*, Graduate School of Business, Stanford University, November 2004.

Palfrey, Thomas R., and Pevnitskaya, Svetlana, "Endogenous Entry and Self-Selection in Private Value Auctions: An Experimental Study," *Unpublished Manuscript*, Department of Economics, University of Southern California, May 2004.

Pevnitskaya, Svetlana, "Endogenous Entry in First-Price Private Value Auctions: the Self-Selection Effect," *Unpublished Manuscript*, Department of Economics, Ohio State University, January 2004

Philipson, Tomas, and Hedges, Larry V., "Subject Evaluation in Social Experiments," *Econometrica*, 66(2), March 1998, 381-408.

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

Rutström, E. Elisabet, "Home-Grown Values and the Design of Incentive Compatible Auctions," *International Journal of Game Theory*, 27(3), 1998, 427-441.

StataCorp, *Stata Statistical Software: Release 8.0* (College Station, TX: Stata Corporation, 2003).

Vella, Francis, "Estimating Models with Sample Selection Bias: A Survey," *Journal of Human Resources*, 33(1), 1998, 127-169.

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