Lost in State Space: Are Preferences Stable?

by

Steffen Andersen, Glenn W. Harrison, Morten I. Lau and E. Elisabet Rutström [*]

May 2007

Working Paper 05-07, Department of Economics,
College of Business Administration, University of Central Florida, 2005

Forthcoming, *International Economic Review*

Abstract. We use field experiments to examine the temporal stability of risk preferences. Stability can mean that a given subject exhibits the same risk attitudes over time, or that their risk attitudes are a stable function of states of nature that change over time. It is quite possible for risk preferences to be stable in either, both or neither of these senses. We report the results of a large-scale panel experiment undertaken in the field designed to examine this issue. Over a 17 month period we elicited risk preferences from subjects chosen to be representative of the adult Danish population. During this period we re-visited many of these subjects and repeated a risk aversion elicitation task. In each visit we also elicited information on their individual characteristics, as well as their expectations about the state of their own economic situation and macroeconomic variables. We control for changes in the subject's perceived states of nature, as well as the possible effects of endogenous sample selection into the re-test. We find some variation in risk attitudes over time, but we do not detect a general tendency for risk attitudes to increase or decrease over a 17-month span. The results also suggest that risk preferences are state contingent with respect to personal finances.

[*] Department of Economics, Copenhagen Business School, Denmark, and Centre for Economic and Business Research, Denmark (Andersen); Department of Economics, College of Business Administration, University of Central Florida, USA (Harrison and Rutström) and Department of Economics and Finance, Durham Business School, Durham University, United Kingdom (Lau). E-mail contacts: SA.ECO@CBS.DK, GHARRISON@RESEARCH.BUS.UCF.EDU, M.I.LAU@DURHAM.AC.UK, and ERUTSTROM@BUS.UCF.EDU. We thank the U.S. National Science Foundation for research support under grants NSF/IIS 9817518 and NSF/HSD 0527675, and the Danish Social Science Research Council for research support under project #24-02-0124. We are grateful to two referees for helpful comments.

Preferences are typically taken as a primitive in economic analysis. When we observe a consumer purchasing different quantities of some good at different points in time, economists presume that the budget constraint has changed in some manner due to price or income changes, asserting preferences to be stationary. The common use of the word "preference" to indicate "choice" can therefore be confusing, since choices can change even though preferences, in the strictest sense of consumer choice theory, are stationary. The power of the assertion that preferences are stable lies in the ability to assign causation between changing opportunity sets and choices in comparative statics exercises. Without such assertions our comparative statics tests would become joint tests of the effect of changing opportunities and changing preferences. To put it differently, all our knowledge about the relationship between opportunities and choices are conditioned on stationarity in preferences.[1]

Stigler and Becker [1977] offer an important perspective on the value of asserting stationarity in preferences. In their opinion the specification of alternative opportunity sets offer more insights than abandoning the assertion of preference stationarity when searching for explanations to choice heterogeneity over time or across individuals. They argue against modeling preferences as changing due to such things as the consumption of "addictive" goods, due to influences of customs and traditions, or fashion and fads, and due to advertising. On the other hand, while maintaining that preferences are stationary, one can alternatively define the arguments of the utility function as state contingent, as long as these states are exogenous to the choices of the agent. This would allow choices to vary with various states of nature. Problems arise, however, when one has to apply this approach empirically. Where does one draw the line in terms of the abstract "states of nature"? The

[1] Asserting that preferences are generally stable is unnecessarily strong. For the purposes of comparative statics, what is required is merely orthogonality between preferences and opportunities such that preferences are stable with respect to the manipulations of the opportunities.

border separating state contingent but stationary preferences from non-stationary ones is not always clearly defined.

We report the results of a large-scale panel experiment undertaken in the field that allows us to examine whether individual risk attitudes are stable over time. In June 2003 we elicited risk preferences from 253 subjects, and the sample was designed to generate a representative sample of the adult Danish population. Between September 2003 and November 2004 we re-visited 97 of these subjects and repeated the tasks. In each visit we also elicited information on their individual characteristics, as well as their beliefs and expectations about the state of their own economic situation and macroeconomic variables. 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 their experimental procedures, but employ subjects that are more representative of individuals affected by public policy changes.[2] We also control for the possible biases of endogenous sample selection that are inherent in virtually any longitudinal experiment that allows subjects to drop out of the panel.

There are several conceptual and measurement issues that arise when one has to determine whether elicited preferences are stable over time. We review these conceptual issues, and the implied metrics, in Section 1. In Section 2 we review our experimental design. We apply extended versions of the basic laboratory procedure designed to elicit more precise responses and control for robustness to framing effects. Section 3 explains the field experiments conducted, with particular attention to the re-sampling procedures. We present the results in Section 4, and draw conclusions in Section 5.

---

[2] Our experiments are "artefactual field experiments" in the terminology of Harrison and List [2004] since we essentially take lab experiments to field subjects.

1. Defining Temporally Stable Preferences

It is possible to define temporal stability of preferences in several different ways, reflecting alternative conceptual definitions and operational measures. Each definition has some validity for different inferential purposes.

A. The Concept of Temporal Stability of Preferences

Temporal stability of risk preferences can mean that subjects exhibit the same risk attitudes over time, or that their risk attitudes are a stable function of states of nature and opportunities that change over time. It is quite possible for risk preferences to be stable in either, both or neither of these senses, depending on the view one adopts regarding the role preference stability takes in the theory. The temporal stability of risk preferences is one component of a broader set of issues that relate to the state-dependent approach to utility analysis.[3] This is a perfectly general approach, where the state of nature could be something as mundane as the weather or as fundamental as the individual's mortality risk. The states could also include the opportunities facing the individual, such as market prices and employment opportunities. Crucial to the approach, however, is the fact that all state realizations must be exogenous, or the model will not be identified.

Problems arise, however, when one has to apply this approach empirically. Where does one draw the line in terms of the abstract "states of nature"? Many alleged violations of expected utility theory amount to claims that a person behaved as if they had one risk preference for one lottery pair and another risk preference for a different lottery pair. Implicit in the claim that these are violations of expected utility theory is the presumption that the differences in the two lottery pairs was not

[3] Hirshleifer and Riley [1992] and Chambers and Quiggin [2000] demonstrate the elegant and powerful representations of decision-making under uncertainty that derive from adopting a state-contingent approach instead of popular alternatives.

some state of nature over which preferences could be different.[4] Similarly, should we deem the preferences elicited with an open-ended auction procedure to be different from those elicited with a binary choice procedure[5] because of some violation of expected utility theory or just some change in the state of nature? Of course, it is a slippery inferential slope that allows "free parameters" to explain any empirical puzzle by shifting preferences. Such efforts have to be guided by direct evidence, lest they become open-ended specification searches.[6]

Proponents of the view that preferences are non-stationary can be found in the discussion about music appreciation in Marshall [1923, p.94], customs and habits in Mill [1872, p.484] and advertising in Galbraith [1958, p.155-56]. A more recent literature (Sunderasan [1989], Constantinides [1990] and Campbell and Cochrane [1999]) argue that habit formation can affect risk preferences over time. This argument relies on a utility function that exhibits temporal non-separability where past consumption habits affect future values.

B. Normalizations

One would expect some variation in observed choices over time, but what measure should be used to say that a certain level of variation is large or small?

One possibility is to examine the average change in risk attitudes, measured on a within-sample basis. That is, we first calculate a measure of how much the risk attitude of a given individual

---

[4] Many of these claims involve evidence from between-sample designs, and rely on the assumption that sample sizes are large enough for randomization to ensure that between-sample differences in preferences (even if they are not state-contingent) are irrelevant. For two careful examples, see Conlisk [1989] and Cubitt, Starmer and Sugden [1998]. There is also a rich literature on the contextual role of extreme lotteries, such that one often observes different behavior for "interior lotteries" that assign positive probability to all prizes as compared to "corner solution lotteries" that assign zero weight to some prizes.

[5] For example, in the famous preference reversals of Grether and Plott [1979].

[6] Stigler and Becker [1977; p.76] note the nature of the impasse: "an explanation of economic phenomena that reaches a difference in tastes between people or times is the terminus of the argument: the problem is abandoned *at this point* to whoever studies and explains tastes (psychologists? anthropologists? phrenologists? sociobiologists?)."

has changed. We then simply ask if the average of these changes is zero. This is a transparent measure, but has the disadvantage that it could be masking large gross changes in risk attitudes that happen to offset each other. The simple correction for this problem is to examine the distribution of changes in risk attitudes, and check that it is unimodal. However, even if this distribution is unimodal, we need some basis for saying that the *variance* of the distribution is small or large. It is possible to decompose the variance and consider whether the variation in risk attitudes over time is less than the variation in risk attitudes across individuals at a given point in time. In the terminology of panel statistics, this is asking if the "within variance" in risk attitudes is less than the "between variance."

We also have to recognize that there may be some attrition issues in any longitudinal setting affecting the stability of the risk attitude estimates. This would arise if subjects condition their participation in later stages of the experiment on unobservable characteristics that correlate with their preferences, or in our case the stability of their preferences. It is not so obvious why one might expect that those subjects with "stable preferences" would self-select into later stages of a panel experiment, but since one might expect some self selection with respect to the preferences themselves, the analysis should account for the possible effects of sample selection.

2. Experimental Design

A. The Basic Elicitation Procedure

We employ a simple experimental measure for risk aversion called a multiple price list (MPL), previously used by Holt and Laury [2002] (HL). Each subject is presented with a choice between two lotteries, which we call A or B. Table 1 illustrates the basic payoff matrix presented to subjects. The first row shows that lottery A offered a 10% chance of receiving 2000DKK and a 90%

chance of receiving 1600DKK. The expected value of this lottery, EV^A, is shown in the fourth-last column as 1640DKK, although the EV columns were not presented to subjects. Similarly, lottery B in the first row has chances of payoffs of 3850DKK and 100DKK, for an expected value of 480DKK. Thus the two lotteries have a relatively large difference in expected values, in this case 1170DKK.[7] As one proceeds down the matrix, the expected value of both lotteries increases, but the expected value of lottery B becomes greater than the expected value of lottery A.

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

We study the effects of experimental conditions in terms of the constant relative risk aversion (CRRA) characterization, employing an interval regression model. The CRRA utility function is defined as $U(y) = (y^{1-r})/(1-r)$, where r is the CRRA coefficient.[8] 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 5 safe choices and then switched to the risky alternatives would have revealed a

---

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

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

CRRA interval between 0.15 and 0.41, and a subject that made 7 safe choices would have revealed a CRRA interval between 0.68 and 0.97, and so on.[9]

B. Extensions

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

First, 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. The four sets of prizes are as follows, in Danish kroner (DKK), with the two prizes for lottery A listed first and the two prizes for lottery B listed next: (A1: 2000, 1600; B1: 3850, 100), (A2: 2250, 1500; B2: 4000, 500), (A3: 2000, 1750; B3: 4000, 150), and (A4: 2500, 1000; B4: 4500, 50). These prizes range from approximately $7.65 to $687. We ask the subject to respond to all four risk aversion tasks and then randomly decide which one to play out.[10] Budget constraints precluded paying all subjects, so each subject is given a 10% chance of actually receiving the payment associated with his decision.

---

[9] To allow for the possibility that relative risk aversion is not constant one can estimate more flexible functional forms, such as the Expo-Power function proposed by Saha [1993]. Maximum likelihood estimates of the Expo-Power model can be used to calculate the relative risk aversion for different income levels. The results documented in Harrison, Lau and Rutström [2007] show that CRRA is an acceptable restriction on the functional form for the sample of the Danish population considered here.

[10] The first set of prizes generates CRRA values at the switch points of -1.71, -0.95, -0.49, -0.14, 0.15, 0.41, 0.68, 0.97 and 1.37; the second set generates values of -1.45, -0.72, -0.25, 0.13, 0.47, 0.80, 1.16, 1.59 and 2.21; the third set generates values of -1.84, -1.101, -0.52, -0.14, 0.17, 0.46, 0.75, 1.07 and 1.51; and the fourth set generates values of -0.75, -0.32, -0.05, 0.16, 0.34, 0.52, 0.70, 0.91 and 1.20. All four sets span 36 distinct CRRA values between -1.84 and 2.21, with roughly 60% of the values reflecting risk aversion. Any scaling of the prizes that is common within a set will preserve the implied CRRA coefficients, so this design could be used with smaller or larger payoffs.

Second, 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. The second extension of the MPL format is to then allow the individual to make choices from refined options within the option last chosen. That is, if someone decides at the first stage to switch from option A to option B between probability values of 0.1 and 0.2, the second stage of an iMPL would then prompt the subject to make more choices *within* this interval, to refine the values elicited.[11]

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 form of one or the other lottery. The iMPL uses the same incentive logic as the MPL. After making all responses, the subject has one row from each stage selected at random by the experimenter.

---

[11] 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.

Finally, we devise a test for framing effects by varying the cardinal scale of the MPL used in the risk aversion task. 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. We develop two asymmetric frames 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). These treatments yield 6 decision rows in the first level of the iMPL, as opposed to the 10 rows in the symmetric frame.[12] As suggested by the treatment names, *skewLO* (*skewHI*) is intended to skew responses to be lower (higher) probabilities if subjects pick in the middle.[13]

C. Panel Experiments

We conducted five series of experiments, beginning in June 2003. Series 1 was the "base camp," where we interviewed 253 subjects. In this experiment we elicited responses for four risk aversion tasks. We also elicited information on individual characteristics and responses to a number of questions about the recent and prospective well-being of the individual, including his perception of the future state of the economy.

Across Series 2 through 5 we re-visited 97 of these 253 subjects. The objective was to re-visit 100 of them, split roughly equally in time. The actual sample sizes were close to this, with 26, 23, 23 and 25 in each of the four series. These follow-up experiments were conducted in September 2003 (3 months after Series 1), November 2003 (5 months after), May 2004 (11 months after), and

---

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

[13] This design does not allow us to reduce responses to a scalar that takes on integer values between 1 and 10 (or 6), and a parametric specification is therefore necessary to compare responses across treatments.

November 2004 (17 months after). In each Series 2 through 5, we elicited responses for the same four risk aversion tasks as in Series 1. The instructions given to subjects are documented in Harrison, Lau, Rutström and Sullivan [2005].

3. Experimental Procedures

A. Sampling Procedures for Series 1

The sampling procedure for the field experiments in Series 1 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, 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. The population represented in these locations amounts to less than 2% of the Danish population, or 493 individuals in our sample of 25,000 from the Civil Registry.
- Third, we assigned each county either 1 session or 2 sessions, in rough proportionality to the population of the county. In total we assigned 20 sessions. Each session consisted of two sub-sessions at the same locale and date, one at 5pm and another at 8pm, and subjects were allowed to choose which sub-session suited them best.
- Fourth, we divided 6 counties into two sub-groups because the distance between some

municipalities in the county and the location of the session would be too large. A weighted random draw was made between the two sub-groups and the location selected, where the weights reflect the relative size of the population in September 2002.

- Fifth, we picked the first 30 or 60 randomly sorted records within each county, depending on the number of sessions allocated to that county. This provided a sub-sample of 600.
- Sixth, we mailed invitations to attend a session to the sub-sample of 600, offering each person a choice of times for the session. Response rates were low in some counties, so another 64 invitations were mailed out in these counties to newly drawn subjects. Everyone that gave a positive response was assigned to a session, and our recruited sample was 268.

Further details of the sampling procedures are documented in Harrison, Lau, Rutström and Sullivan [2005].

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.[14] These experiments were conducted in June of 2003, and a total of 253 subjects participated in the experiments.[15] 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.

---

[14] 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.

[15] 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.

B. Sampling Procedures for Series 2 Through 5

Between September 2003 and November 2004 we re-visited 97 of the 253 subjects who participated in Series 1 and repeated the four risk aversion tasks. Each subject was interviewed in private in the four new series of experiments, because attendance at the experiments otherwise would have been too low. To minimize travel times for subjects and encourage higher attendance, we offered to conduct the experiment at their private residence, or at another convenient location of their choice (e.g., the hotel where the Series 1 experiment took place). There were four re-visits, which we call Series 2 through 5.

There were four steps in the construction of this sample. First, we assigned each of the 14 counties either 1 or 2 numbers, in rough proportionality to the population of the county. In total we assigned 20 numbers that also reflect the distribution of the 20 sessions in the sampling procedures for Series 1. Second, although Denmark is a relatively small country, it was again necessary to consider logistical constraints, and we randomly picked 4 of the 20 numbers for each of Series 2 through 5. This procedure implied that we would re-visit subjects from at most four counties in each new series of experiments. Third, we picked the first 25% or 50% of the randomly sorted records within each county, depending on the one or two numbers allocated to each county. This provided a sub-sample of 100 subjects. Fourth, we contacted subjects by phone and invited them to participate again in the experiments. It was difficult to get in contact with some subjects, and other subjects did not want to participate again in the experiments, so we invited the next subjects on the randomly sorted list of former participants until 26 people were signed up for each new series of experiments.[16]

[16] We anticipated that a few subjects would cancel the meeting and thus signed up one extra subject for each new series. We were unable to get in contact with 15% of the subjects, and only 16% of the subjects refused or were not able to participate again.

The experiments were again computerized, and all the experimental procedures, tasks, frames and monetary incentives were similar to those used in Series 1 and described in Harrison, Lau, Rutström and Sullivan [2005].[17]

D. The Sample

Table 2 provides the definitions of the explanatory variables used in the statistical analysis 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. We obtained the same share of men and women as the Danish population, although we observe relatively more subjects between 40 and 50 years of age in our two samples compared to the population.[18] There are no significant differences in the composition of the two samples, although there were relatively more subjects with substantial higher education and higher level income who participated in Series 2-5 of the experiments compared to Series 1.

The experimental design allows us to correct for sample selection into the first and later series of the experiments. There is little information on the subjects who declined our invitation to participate in the first series. We therefore focus on possible attrition bias and take the sample of 253 subjects in Series 1 of the experiments as the population, and use the elicited information on individual characteristics to control for sample selection into the re-tests. Harrison, Lau and

[17] Since the subjects had prior experience with the experimental procedures, we did not spend time training subjects on the iMPL and randomization procedures. However, we maintained the trainer task in which payments were in the form of candies.

[18] The population shares are based on information from the household expenditure survey conducted by Statistics Denmark. Households in the expenditure survey are randomly selected by their address and classified according to the characteristics of the person with the highest income in the household. Subjects in our experiments were randomly selected by their personal ID number and classified according to the characteristics of the subject. Although we applied the same household classifications as Statistics Denmark in our socio-demographic questionnaire, it is not possible to make direct comparisons between our two samples and the information from the household expenditure survey.

Rutström [2007] use data from Series 1 and correct for sample selection in their analysis of risk attitudes. They find evidence of sample selection into Series 1, but the estimates of individual characteristics are similar across specifications with and without correction for sample selection. It is also possible that our use of show-up fees may generate a sample that was more or less risk averse than would otherwise have been observed. The initial recruitment letter clearly stated that subjects would be paid 500DKK for their participation and that there would be some randomization involved in determining earnings during the experiment. However, we did not vary the fixed or random payments between subjects and are not able to analyze this issue further without additional information.

4. Results

A. Raw Results

Figure 1 collects the raw results of our experiments. It displays the distribution of elicited relative risk aversion by horizon and risk aversion task, using box plots for each task.[19] Panel A of Figure 1 shows the raw results of the Series 1 experiments, and Panel B illustrates the raw results of the Series 2 through 5 experiments. The raw data here are the mid-points of the elicited CRRA interval, after all iterations of the iMPL have been completed.[20] These comparisons are unconditional, in the sense that they do not control for changes in states of nature or possible attrition.

We observe variations of elicited CRRA across subjects, with a mean of 0.62, a median of 0.64 and a standard deviation of 0.80. This distribution is consistent with comparable estimates

[19] A box plot shows the median as a solid dot, the inter-quartile range as a shaded rectangle, and the range in the outer "whiskers." The interquartile range is the 25th and 75th percentile.

[20] The distribution of the elicited CRRA interval size is right skewed with a mean of 0.164 and a median of 0.031. More than 25% of the observations are point estimates and the inter-quartile range is between 0 and 0.075.

obtained in the United States, using college students and an MPL design, by Holt and Laury [2002] and Harrison, Johnson, McInnes and Rutström [2005a][2005b]. Few subjects are risk loving or risk neutral. Risk aversion is by far a better characterization of the risk preferences of the average Dane.

The raw results indicate that there is considerable variation in CRRA across the sample. The inter-quartile range for all observations is between 0.17 and 1.07. Of course, this variation could just reflect heterogeneity in subjective preferences or differences in risk attitudes across task or series. Pooling over all tasks and series we find a between-subject standard deviation of 0.69 which is greater than the within-subject standard deviation of 0.52. The latter is calculated based on eight tasks for each subjects, four in each of the two series they participated in. This shows that the variation in responses due to time or task is smaller than the variation due to individual heterogeneity in subjective preferences.[21] We also find a within-series standard deviation of 0.79 which is higher than the between-series standard deviation of 0.22. This indicates that the variation in risk attitudes across time is smaller than the variation across subjects and tasks.

We can also look at the correlation between CRRA values elicited at different points in time and estimate four linear regression models, one model for each of the four risk aversion tasks. The dependent variable is the elicited CRRA in Series 2-5 and the independent variable is elicited CRRA in Series 1. The results show that there is a significant correlation between the elicited CRRA values in Series 2-5 and those values elicited in Series 1. The coefficient is significantly different from zero in all four models and is equal to 0.43 (*p*-value of 0.033) for task 1 comparisons, 0.46 (*p*-value of 0.002) for task 2, 0.58 (*p*-value of less than 0.001) for task 3 and 0.34 (*p*-value of 0.033) for task 4 comparisons. However, the four coefficients are significantly different from one, which indicates

[21] We can look at the data in further detail and consider the variation over time for each pair of tasks that the design allows us to compare. The between-subject standard deviation is considerably higher than the within-subject standard deviation for all four paired tasks. The between-subject (within-subject) standard deviation is 0.73 (0.33) for task 1; 0.99 (0.33) for task 2; 0.80 (0.27) for task 3; and 0.52 (0.23) for task 4.

that there is some variation in risk attitudes over time.

There are limits to what one can infer from these raw, unconditional data and between-sample comparisons. For example, within the inter-quartile range of all the observations presented in Figure 1 there could be some subjects with sharply increasing or declining relative risk aversion over time, and others that are simply high-variance over time. To better evaluate the hypotheses of interest we must turn to within-sample comparisons across time.

Figure 2 displays the data for the within-sample comparisons that our experimental design was constructed to allow. The design allows us to generate four observations for each subject: one observation for each of the four comparable tasks at two different points in time. The values show the difference between the midpoint of the elicited CRRA interval from the same subject and task at two different points in time.[22] If there was no change in that elicited CRRA then the data point underlying the histogram in Figure 2 would be zero. If the CRRA had increased in the later series, the data point would have been positive. The differences in relative risk aversion illustrated in Figure 2 do not have any apparent tendency to be positive or negative. The mean of the distribution is −0.07, while the median is 0.00 and the inter-quartile range is between -0.49 and 0.41. The Figure is clearly uni-modal so that the zero median is not masking the presence of large segments of subjects with positive and negative changes.

We can go one step further and classify subjects according to four groups: (i) subjects with no change in risk attitudes at two different points in time, (ii) subjects with higher relative risk aversion at a later date, (ii) subjects with lower relative risk aversion at a later date, and (iv) subjects who reveal both higher and lower relative risk aversion over time. Figure 3 displays the shares of subjects that fall into those four categories conditional on a noise parameter. The noise parameter

[22] We report 287 observations. We treat open-ended CRRA intervals as missing values and do not impose an upper or a lower bound on these values.

defines a symmetric, uniform interval around the elicited difference in relative risk aversion at two different points in time. When the noise parameter is equal to zero, we find that 21% of the subjects are more risk averse across all four comparable tasks at the later date, 27% of the subjects are less risk averse over time, and 52% of the subjects are more and less risk averse across the four comparable tasks. The share of subjects with no difference in relative risk aversion increases with the size of the noise parameter by definition, while the share of subjects with positive or negative differences in relative risk aversion falls to zero when the level of noise is sufficiently large; which in this case is 2.0 on either side of the elicited difference in relative risk aversion.

The unconditional data shows that there is no general positive or negative bias in elicited relative risk aversion over time. A direct test of the hypothesis that individual risk preferences are constant over time is possible by estimating a model in which the difference in relative risk aversion for any two comparable tasks over time is the dependent variable. Since the dependent variable is a within-sample comparison, our design provides controls for the characteristics of the individual.[23] However, we also need to control for possible changes in the states of nature facing our subjects over the time frame between our experiments, as well as attrition.

B. Controlling for Changes in States of Nature

Although we visited the same person at a later date, the fact that a certain amount of time had to pass means that there may have been changes in the states of nature that could have an effect on elicited risk attitudes. Common sense and interpretation suggest that individual risk preferences could be state-dependent, and we evaluate the extent to which observable changes in states of nature

[23] We repeated the demographic questions in the re-tests, and there were no significant changes in those observed demographics.

can explain the observed changes in elicited risk aversion over time.[24]

In each series we asked subjects to respond to seven questions about their perception of the state of the economy in general and their own personal financial situation. These questions are similar to those asked in the Survey of Consumers that is conducted regularly by the University of Michigan (http://www.sca.isr.umich.edu/). The survey of Consumers contains approximately 50 questions and we chose 7 of these as indicators of states of nature. In each case we asked for their perception for horizons, denoted by X below, of 1, 4, 6, 12, 18 and 24 months:

- Would you say that you and your family are better off or worse off financially than you were X months ago?
- Now looking ahead, do you expect any major change in your family situation that will lead to higher expenses or lower expenses during the next X months?
- Do you expect any major change in your family situation that will lead to higher earnings or lower earnings during the next X months?
- On balance, do you think that you and your family will be better off or worse off financially X months from now?
- Turning to the economic conditions in the country as a whole, would you say that at the present time economic conditions are better or worse than they were X months ago?
- Do you think that there will be more or less unemployment during the next X months?
- Do you think that interest rates for borrowing money will go up or go down during the next

[24] Horowitz [1992] and Harrison, Johnson, McInnes and Rutström [2005a] demonstrate the temporal stability of risk attitudes in lab experiments over a period of up to 4 months. Horowitz [1992; p.177] collects information on financial characteristics of the individual to control for changes in state of nature, but does not report if it changed the statistical inference about temporal stability. Harrison, Johnson, McInnes and Rutström [2005a] consider the temporal stability of risk attitudes in college students over a four week period, and do not control for changes in state of nature. Our experiments extend their design in several ways. We use a much longer time span, control for changes in state of nature, and use a stratified sample of a broader population.

X months?

We readily concede that these questions do not exhaust the set of conceivable events that could occur over the horizon of interest,[25] but they are certainly a good general place to start looking for possible effects from changes in states of nature.

We construct a variable for each subject using their responses to these questions. For each question we asked if they thought that there would be an improvement, a worsening, no change, or whether they did not know. We coded improvements as 1, a worsening as -1, no change as 0, and don't know as missing. For each question and horizon we can then calculate the change from the response in the Series 1 session to the comparable response in the later session. Thus for any question-horizon-subject combination the difference could be +2, +1, 0, -1, -2 or missing. We use state of nature indicators for one of the horizons, and that horizon depends on the time span between the two experiments. For Series 2 observations we apply the state of nature questions for a horizon of 4 months, for Series 3 observations we use questions for a horizon of 6 months, and so on. This value reflects the extent to which they expect an improvement or worsening in each state over the given horizon. There are many ways to summarize these data, but this statistic seems sensible here.

Table 3 describes these statistics of the changes in states of nature. A positive value means that subjects believe that the specific state of nature is improved at the time of the later sessions compared to Series 1, and vice versa for negative values. Subjects are on average marginally more optimistic about their own personal financial situation at the later session compared to the base camp, although they seem to have become a bit more pessimistic about future personal income and expenditures. The subjects have a more positive impression of the current state of the economy

[25] In the demographic survey conducted in Harrison, Lau and Williams [2002] we literally had one subject ask us if he wanted us to state his current sex or the sex he would be at the end of the longest horizon.

compared to the past, and they have become more optimistic about the general level of unemployment. Finally, we observe that subjects have become more pessimistic about changes in interest rates and tend to believe that the interest rate for borrowing money will go up in the near future.

C. Statistical Controls for Paired Risk Aversion Tasks

Table 4 presents the maximum likelihood estimates of two linear regression models of differences in elicited CRRA values from the same subject. The dependent variable is the within-sample difference in CRRA for the "same task" at different times. The controls include a constant, which in this case shows us the average change in CRRA. Additional controls include the state of nature indicators described above. We allow for the complex survey design and adjust for the fact that subjects in one county were selected independently of subjects in other counties, that sample weights for each subject reflect the adult population of Denmark, as well as the possibility of correlation between responses from the same subject.[26]

The two models illustrate the effects of correcting for attrition bias between the first and later series. The first model in panel A of Table 4 does not correct for attrition bias while the second model in panel B does. If we take the sample of 253 subjects in Series 1 as the population for present purposes, the sub-sample of 97 subjects in later series can be viewed as the result of

[26] 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]). Wooldridge [2003] reviews some issues in the use of clustering for panel effects, in particular noting that significant inferential problems may arise with small numbers of panels.

exogenous and endogenous selection processes. The exogenous processes were the ones we used to invite subjects, described earlier. No attempt was made to identify subjects to recruit for the follow-up experiments based on their initial responses. However, we could not control who would accept our invitation to participate again, and therein lies the potential for sample selection through attrition.[27]

The attrition equation in panel B2 of Table 4 reflects maximum-likelihood estimation of a standard sample selection model.[28] The equation is essentially a probit model of participation in the re-test, and uses a specification that also allows for the complex survey design. We include measures of earnings in the prior session, a dummy for the experimenter in the prior session, regional dummies, and individual characteristics.[29]

Temporal stability of risk attitudes is measured by the constant in the estimated equation of the difference in CRRA. Panel A of Table 4 shows that the coefficient is -0.06 in the model that does not correct for attrition bias, indicating a small decrease in average CRRA over time. This estimate has a *p*-value of 0.38 and is not significantly different from zero, so we cannot reject the hypothesis that risk preferences are constant over time. We get a similar result when the model corrects for attrition bias. Panel B shows that the constant in the main CRRA equation has a coefficient of 0.09 and a *p*-value of 0.73. The ancillary parameter ρ measures the correlation between

[27] The standard methods of sample selection apply in the case of sample attrition, as noted by Heckman [1979; p.154]. If we were analysing behaviour in each of the series as a panel, rather than the difference between responses in the two series that a given subject participated in, one could use comparable methods developed by Hausman and Wise [1979].

[28] 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.

[29] Earnings in the prior session is represented by two binary variables. One variable indicates earned income in prior risk aversion tasks, and the other variable indicates earned income in prior individual discount rate tasks. These discount rate tasks were presented to subjects after the risk aversion tasks, and each subject had a 10% chance of winning at least 3,000DKK. We consider the characteristics of subjects at the time that they were recruited for the Series 1 experiments. There are very few changes in these characteristics between the two sessions.

the residuals of the attrition equation and the CRRA equation. This coefficient is −0.42, has a standard error of 0.63, and has a 95% confidence interval with values of −1.45 and 0.78. Thus we see no evidence of attrition bias on the main estimates of differences in CRRA over time.

It is possible to break up the constant in Table 4 into a series of binary dummies for each of the re-test comparisons. For each of the four re-test sessions we have 4 separate tasks for each subject, implying 16 possible effects. Table 5 shows the estimates for these variables, excluding all of the estimates for the other variables from Table 4 although they were controlled for in the estimation. Looking at the model with no correction for attrition we do not find a particular positive or negative bias in the coefficients of the 16 binary variables. There is relatively little variation in risk attitudes over time and only 3 of the 16 coefficients are significantly different from zero: the coefficient of task 2 in Series 2 is equal to 0.57 and has a *p*-value of less than 0.001, while the coefficient of task 1 and 4 in Series 3 is −0.36 (*p*-value of 0.03) and −0.41 (*p*-value of 0.04), respectively.[30]

Turning to the model that corrects for attrition, we find that 14 of the 16 coefficients are positive. Three of these positive coefficients (task 2, 3 and 4 in Series 2) are significantly different from zero at the 5% level. These estimates suggest that subjects are more risk averse in Series 2, which is three months after the first series of experiments. However, we do not find significant differences in average CRRA (at the 5% level) in later Series 3-5, although they seem to become more risk averse with respect to task 4 in Series 4 and 5. The results do not show evidence of attrition bias that would significantly influence the main estimates of differences in CRRA over time

[30] We have also estimated the model using the mean of the four observations per subject. We thus collapse the four observations per subject to one observation. The model includes four independent binary variables: one binary variable for each of the later series. We find that the binary variable for Series 3 has a coefficient of −0.25 and has a *p*-value of 0.036. The other three binary variables are not significant at the 10% confidence level. This result suggests that subjects are less risk averse six months after the first set of experiments.

consistently across tasks. The estimates in the two models thus suggest that there is some variation in risk attitudes over time, but with no particular negative or positive bias.[31]

Returning to Table 4, we find that two of the state of nature indicators, Ds_fin0 and Ds_exp, are negative and significantly different from zero using both models. These two variables measure how positive subjects are about their own financial situation compared to the past and their optimism about their future personal expenditures. The state of nature indicators that relate to the general economy have very small coefficients and there is no significant correlation with changes in risk attitudes over time. Thus we see that risk preferences are state dependent with respect to personal finances while the overall economic climate seems to have little impact on individual risk attitudes.

Finally, we can use an alternative way of breaking up the constant that measures the average change in CRRA over time. We repeat the statistical analysis in Table 4 and include one demographic characteristic (listed in Table 2) at a time to get a measure of individual heterogeneity in behavior over time. Compared to incorporating all demographic characteristics simultaneously, this method allows the coefficient to capture the joint influence on behavior from all characteristics that are correlated with the included characteristic.[32] We find that the variation in risk attitudes over time is not significantly correlated with any of the demographic characteristics. None of the coefficients that represent individual characteristics are significantly different from zero, even at the 10% confidence level, implying that we cannot identify any demographic group of subjects with unstable temporal preferences.

---

[31] These conclusions also hold if the dependent variable measures the difference between the endpoints, and not the midpoints, of the elicited CRRA intervals. The elicited CRRA intervals are small in size and the results are not sensitive to alternative interval estimators.

[32] A discussion of this estimation method, and the manner in which it complements the standard method in which all characteristics are jointly controlled for, is provided by Harrison, Lau and Williams [2002] and Harrison, Lau and Rutström [2007].

6. Conclusions

We find some variation in risk attitudes over time, but we do not detect a general tendency for risk attitudes to increase or decrease over a 17-month span. This is true whether or not we correct for possible biases due to endogenous attrition. We are able to identify two states of nature which appear to be associated with changes in risk attitudes: as subjects become more positive about their current finances and more optimistic about future expenditures, risk attitudes tend to decline. The results thus suggest that risk preferences are state contingent with respect to personal finances. We do not find any demographic group of subjects for whom risk preferences change over time in a significant way.

Our approach has the strength of directly eliciting the preferences we are interested in using controlled methods. It has obvious limitations of application. The 17-month horizon we compared preferences over could be extended. At the risk of encouraging more severe sample selection problems, one could try to recruit sub-samples that had experienced some threshold changes in their states of nature, such as unemployment, divorce or childbirth.[33] Finally, the stakes could be higher although we use large stakes compared to other studies. Such extensions are straightforward.

[33] Minimally invasive survey procedures, such as telephone calls, could be used to screen subjects for more extensive re-tests.

Table 1: Typical Payoff Table for Risk Aversion Experiments

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

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

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

Table 2: List of Demographic Characteristics and Descriptive Statistics

Variable	Definition	Series 1 Sample Mean	Series 2-5 Sample Mean	Danish Population
female	Female	0.51	0.53	0.50
young	Aged less than 30	0.17	0.14	0.19
middle	Aged between 40 and 50	0.28	0.29	0.20
old	Aged over 50	0.37	0.39	0.39
single	Lives alone	0.20	0.16	0.40
kids	Has children	0.28	0.32	0.24
hhd_2	2 people in the household	0.42	0.46	0.32
hhd_3	3 people in the household	0.17	0.20	0.12
hhd_4	4 people or more in the household	0.22	0.20	0.16
owner	Owns own home or apartment	0.69	0.67	0.51
retired	Retired	0.16	0.17	0.28
student	Student	0.09	0.09	0.02
skilled	Some post-secondary education	0.38	0.31	0.38
longedu	Substantial higher education	0.36	0.47	0.18
IncLow	Lower level income	0.34	0.31	0.43
IncHigh	Higher level income	0.33	0.40	0.34
	Number of subjects	253	97	

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.

Note: The third column presents the Danish population shares in 2003. Statistics on gender and age are based on information from the Danish Civil Registration Office. We include people between 19 and 75 years of age in the comparisons. Statistics on the other classifications are based on information from the household expenditure survey conducted by Statistics Denmark. A private household is defined as an economic unit, i.e. a group of people living together that share income and expenses, and the households are classified according to the person with the highest income in the household. The data is collected from approximately 3,000 households over a three year period from 2002 to 2004.

Figure 1: Distribution of Risk Attitudes

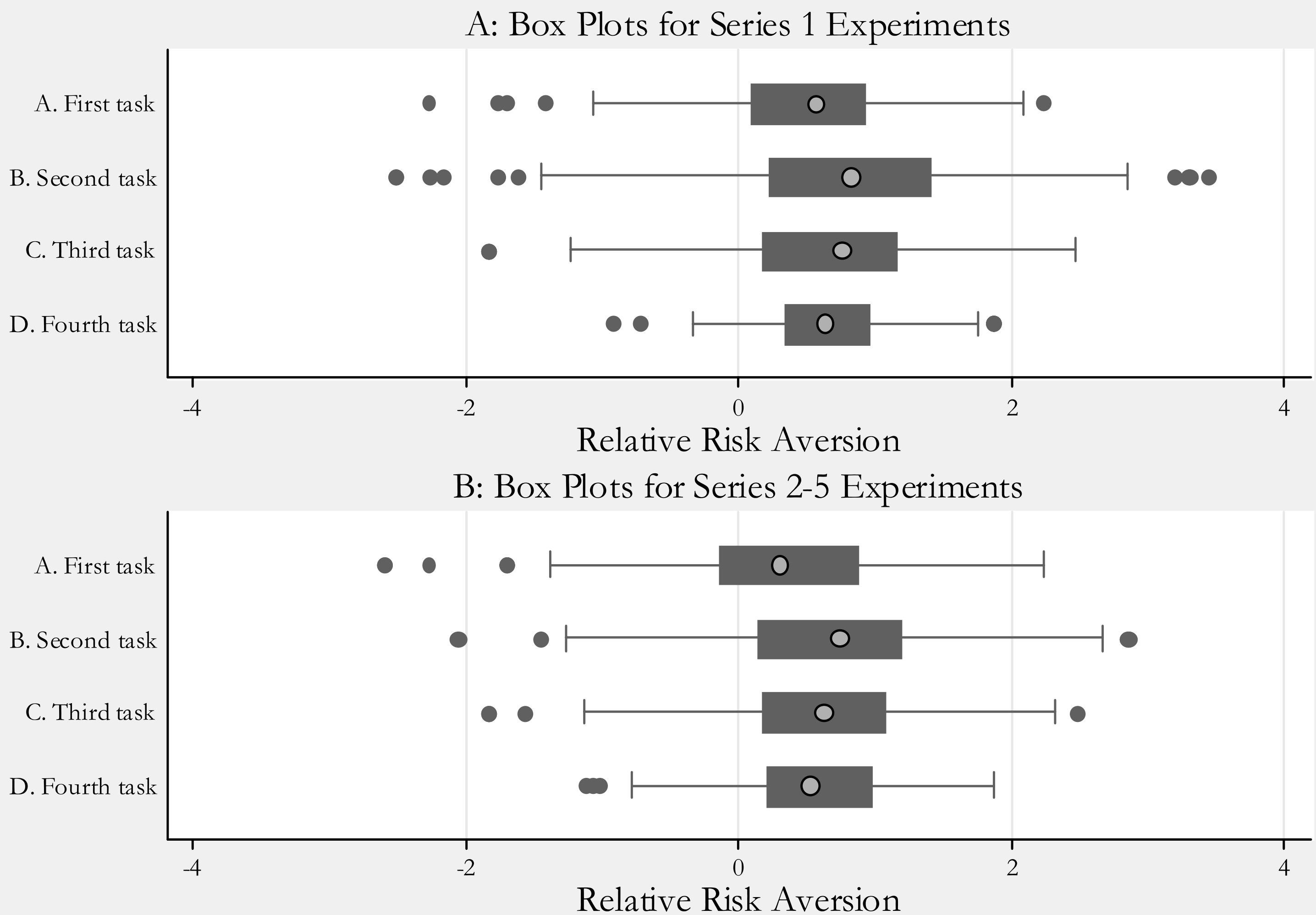
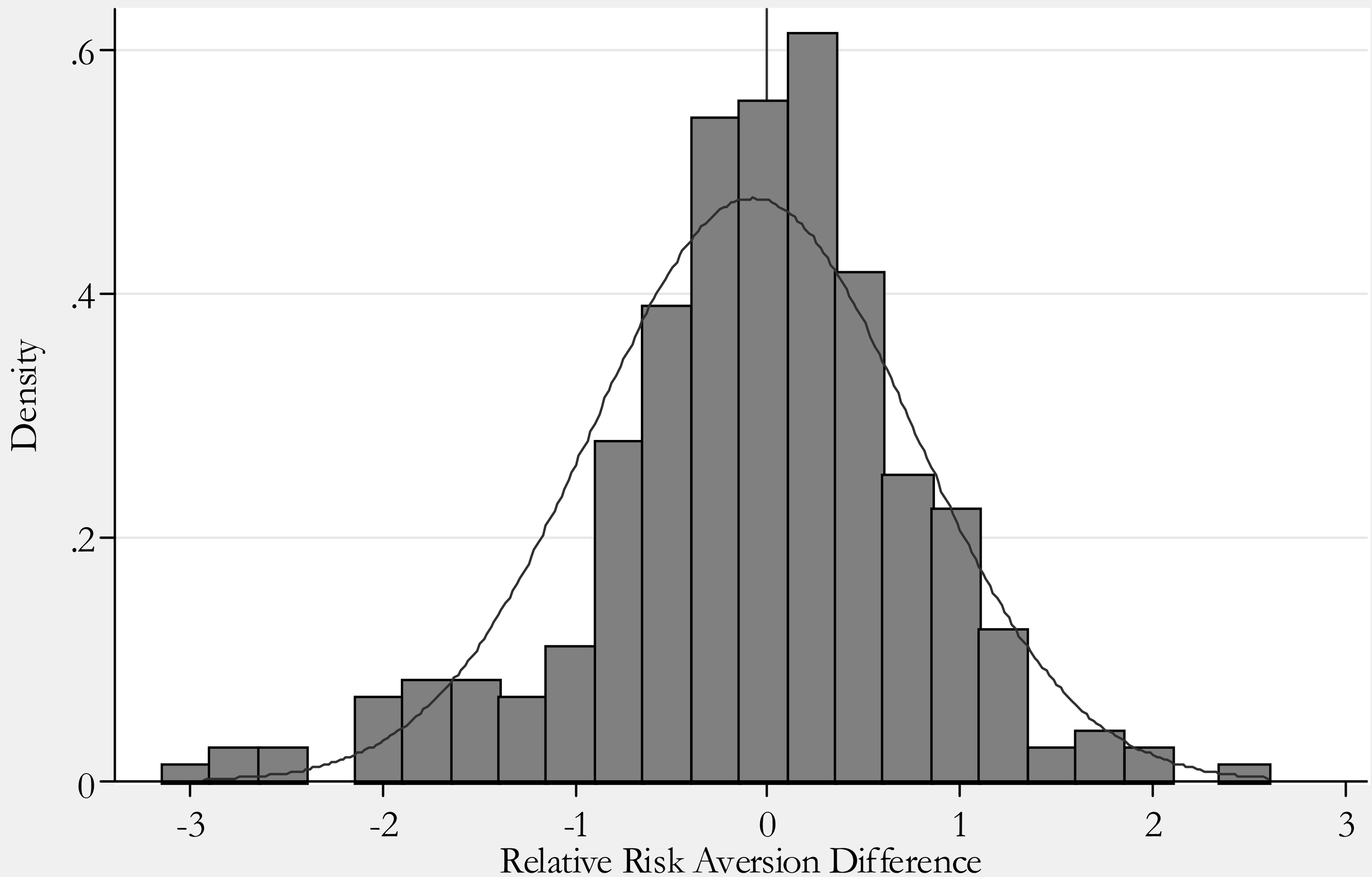


Figure 2: Within-Subject Differences
in Elicited Relative Risk Aversion



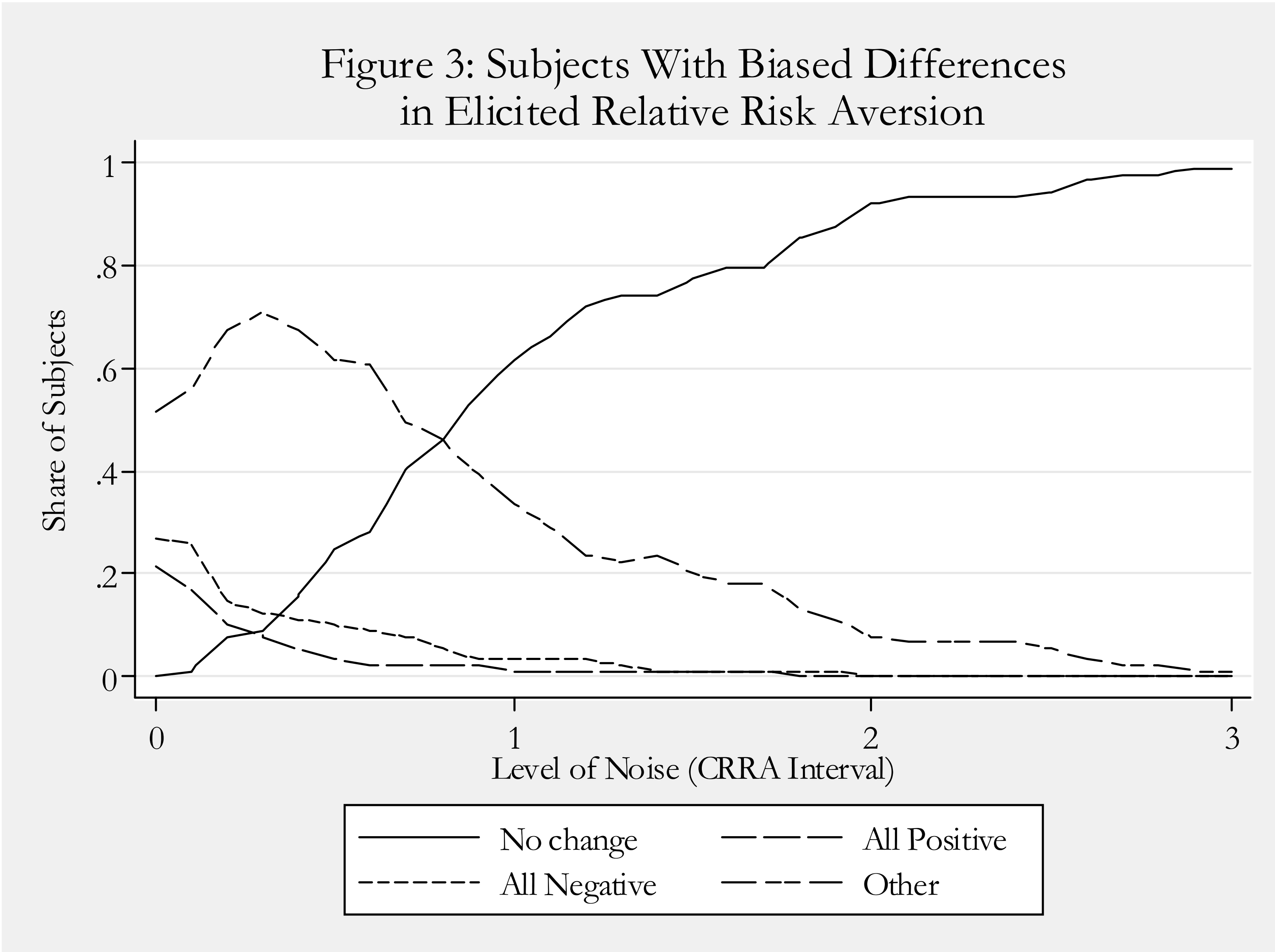


Figure 3: Subjects With Biased Differences
in Elicited Relative Risk Aversion
Share of Subjects
0
.2
.4
.6
.8
1
0
1
2
3
Level of Noise (CRRA Interval)
No change
All Positive
All Negative
Other

Table 3: Descriptive Statistics for the Changes in States of Nature

N=388, based on 97 subjects

Variable	Description	Mean	Standard Deviation	Min	Max
Ds_fin0	Financial situation compared to the past	0.02	0.7	-2	2
Ds_exp	Future personal expenditures	-0.07	0.6	-2	2
Ds_inc	Future personal earnings	-0.07	0.48	-1	1
Ds_fin1	Future financial situation	0.02	0.61	-1	2
Ds_eco	Economic situation compared to the past	0.12	0.75	-2	2
Ds_emp	Future general employment level	0.08	0.71	-2	2
Ds_int	Future general interest rate level	-0.22	0.71	-2	1

Legend: Variable "ds_fin0" indicates if the subject is better off or worse off financially than he were X months ago, variable "ds_exp" indicates if the subject expects any major change that will lead to higher expenses or lower expenses during the next X months, variable "ds_inc" indicates if the subject expects any major change that will lead to higher earnings or lower earnings during the next X months, variable "ds_fin1" indicates if the subject thinks that he will be better off or worse off financially X months from now, variable "ds_eco" indicates if the subject would say that at the present time economic conditions are better or worse than they were X months ago, variable "ds_emp" indicates if the subject thinks there will be more or less employment during the next X months, and variable "ds_int" indicates if the subject thinks that interest rates for borrowing money will go up or down during the next X months. "X" months refer to 1 month, 4 months, 6 months, 12 months, 18 months and 24 months.

Table 4: Within-Sample Differences in Elicited Risk Attitudes

Survey linear regression model of difference in elicited CRRA (287 observations)

Variable	Description	Estimate	Standard Error	*p*-value	Lower 95% Confidence Interval	Upper 95% Confidence Interval
			A. Without correction for attrition			
Constant		-0.06	0.07	0.38	-0.19	0.07
Ds_fin0	Financial situation compared to the past	-0.23	0.09	0.01	-0.40	-0.05
Ds_exp	Future personal expenditures	-0.18	0.09	0.05	-0.35	0.00
Ds_inc	Future personal earnings	-0.06	0.14	0.68	-0.34	0.22
Ds_fin1	Future financial situation	0.17	0.11	0.12	-0.04	0.39
Ds_eco	Economic situation compared to the past	-0.03	0.07	0.68	-0.18	0.12
Ds_emp	Future general employment level	-0.01	0.07	0.86	-0.16	0.13
Ds_int	Future general interest rate level	0.00	0.08	0.99	-0.16	0.17
			B. With correction for attrition			
Constant		0.09	0.26	0.73	-0.43	0.61
Ds_fin0	Financial situation compared to the past	-0.21	0.10	0.04	-0.41	-0.01
Ds_exp	Future personal expenditures	-0.21	0.10	0.05	-0.42	0.00
Ds_inc	Future personal earnings	-0.05	0.14	0.75	-0.33	0.24
Ds_fin1	Future financial situation	0.16	0.12	0.18	-0.07	0.39
Ds_eco	Economic situation compared to the past	-0.06	0.09	0.56	-0.24	0.13
Ds_emp	Future general employment level	-0.04	0.09	0.67	-0.21	0.14
Ds_int	Future general interest rate level	0.02	0.09	0.86	-0.16	0.19
			B2. Attrition equation			
Constant		1.39	1.07	0.20	-0.73	3.52
Rearned	Earned income in prior RA task	0.31	0.38	0.41	-0.43	1.06
IDRearned	Earned income in prior IDR task	0.62	0.57	0.28	-0.51	1.76
PriorExp	Prior experimenter dummy	-0.02	0.42	0.97	-0.84	0.81
Region_2	Rest of Zealand and Funen	0.65	0.43	0.13	-0.20	1.49
Region_3	Jutland	-0.03	0.43	0.94	-0.89	0.82
female	Female	-0.59	0.39	0.14	-1.36	0.19
young	Aged less than 30	-0.03	0.58	0.96	-1.17	1.11
middle	Aged between 40 and 50	0.09	0.49	0.86	-0.88	1.05
old	Aged over 50	0.41	0.53	0.45	-0.65	1.46
single	Lives alone	-0.77	0.69	0.27	-2.14	0.60
kids	Has children	-0.36	0.48	0.46	-1.30	0.59
hhd_2	2 people in household	-0.81	0.80	0.32	-2.39	0.78
hhd_3	3 people in household	0.09	0.72	0.90	-1.34	1.53

hhd_4	4 or more people in household	0.38	0.87	0.66	-1.34	2.09
owner	Owns own home or apartment	-1.02	0.48	0.04	-1.97	-0.07
retired	Retired	0.23	0.54	0.68	-0.85	1.30
student	Student	-0.16	0.57	0.77	-1.29	0.96
skilled	Some post-secondary education	-0.25	0.45	0.58	-1.14	0.65
longedu	Substantial higher education	0.32	0.47	0.49	-0.61	1.25
IncLow	Lower level income	0.81	0.45	0.07	-0.07	1.70
IncHigh	Higher level income	0.11	0.39	0.77	-0.65	0.88
rho	Error correlation	-0.42	0.63		-1.45	0.78

Table 5: Decomposition of Within-Sample Differences in Elicited Risk Attitudes

Survey linear regression model of difference in elicited CRRA (287 observations)

		Without correction for attrition			With correction for attrition		
Variable	Description	Estimate	Standard Error	*p*-value	Estimate	Standard Error	*p*-value
RApair_2_1	Series 2, task 1	-0.37	0.26	0.15	0.07	0.34	0.83
RApair_2_2	Series 2, task 2	0.57	0.18	0.00	0.97	0.34	0.01
RApair_2_3	Series 2, task 3	0.22	0.19	0.26	0.61	0.31	0.05
RApair_2_4	Series 2, task 4	0.13	0.11	0.25	0.44	0.20	0.03
RApair_3_1	Series 3, task 1	-0.36	0.17	0.03	-0.06	0.23	0.80
RApair_3_2	Series 3, task 2	-0.05	0.16	0.77	0.24	0.24	0.32
RApair_3_3	Series 3, task 3	-0.15	0.23	0.51	0.13	0.26	0.61
RApair_3_4	Series 3, task 4	-0.41	0.20	0.04	-0.16	0.21	0.45
RApair_4_1	Series 4, task 1	0.12	0.25	0.63	0.38	0.29	0.19
RApair_4_2	Series 4, task 2	-0.07	0.23	0.77	0.21	0.31	0.49
RApair_4_3	Series 4, task 3	-0.10	0.14	0.47	0.10	0.18	0.57
RApair_4_4	Series 4, task 4	0.42	0.30	0.16	0.67	0.38	0.08
RApair_5_1	Series 5, task 1	-0.12	0.19	0.53	0.10	0.18	0.59
RApair_5_2	Series 5, task 2	-0.26	0.18	0.16	0.00	0.26	1.00
RApair_5_3	Series 5, task 3	-0.15	0.17	0.38	0.10	0.25	0.68
RApair_5_4	Series 5, task 4	0.03	0.09	0.73	0.21	0.11	0.07

References

Campbell, John Y., and Cochrane, John H., "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107, 1999, 205-251.

Chambers, Robert G., and Quiggin, John, *Uncertainty, Production, Choice, and Agency: The State-Contingent Approach* (New York, NY: Cambridge University Press, 2000).

Conlisk, John, "Three Variants on the Allais Example," *American Economic Review*, 79(3), June 1989, 392-407.

Constantinides, George M., "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy*, 98, 1990, 519-543.

Cubitt, Robin P.; Starmer, Chris, and Sugden, Robert, "Dynamic Choice and the Common Ratio Effect: An Experimental Investigation," *Economic Journal*, 108, September 1998, 1362-1380.

Galbraith, John K., *The Affluent Society* (Boston: Houghton Mifflin, 1958)

Grether, David M., and Plott, Charles R., "Economic Theory of Choice and the Preference Reversal Phenomenon," *American Economic Review*, 69(4), September 1979, 623-648.

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

Harrison, Glenn W.; Johnson, Eric; McInnes, Melayne M., and Rutström, E. Elisabet, "Risk Aversion and Incentive Effects: Comment," *American Economic Review*, 95(3), June 2005b, 897-901.

Harrison, Glenn W.; Lau, Morten I., and Rutström, E. Elisabet, "Estimating Risk Attitudes in Denmark: A Field Experiment," *Scandinavian Journal of Economics*, 109, 2007 forthcoming.

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

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

Hausman, Jerry A., and Wise, David A., "Attrition Bias in Experimental and Panel Data: The Gary

Income Maintenance Experiment," *Econometrica*, 47(2), March 1979, 455-473.

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.

Hirshleifer, Jack, and Riley, John G., *The Analytics of Uncertainty and Information* (New York, NY: Cambridge University Press, 1992).

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

Horowitz, John K., "A Test of Intertemporal Consistency," *Journal of Economic Behavior and Organization*, 17, 1992, 171-182.

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

Marshall, Alfred, *Principles of Economics* (London: Macmillan, 8th Edition, 1923).

Mill, John S., *A System of Logic* (London: Longmans, Green & Co., 8th Edition, 1872).

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

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

Stigler, George J., and Becker, Gary S., "De Gustibus Non Est Disputandum," *American Economic Review*, 67(2), March 1977, 76-90.

Sundaresan, Suresh M., "Intertemporally Dependent Preferences and the Volatiliuty of Consumption and Wealth," *Review of Financial Studies*, 2, 1989, 73-88.

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.

Wooldridge, Jeffrey, "Cluster-Sample Methods in Applied Econometrics," *American Economic Review (Papers & Proceedings)*, 93, May 2003, 133-138.