Elicitation Using Multiple Price List Formats

by

Steffen Andersen, Glenn W. Harrison, Morten Igel Lau and E. Elisabet Rutström†

October 31, 2004

Abstract. We examine the properties of a popular method for eliciting choices and values from experimental subjects, the multiple price list format. The main advantage of this format is that it is relatively transparent to subjects and provides simple incentives for truthful revelation. The main disadvantages are that it only elicits interval responses, and could be susceptible to framing effects. We consider extensions to address and evaluate these concerns. We conclude that although there are framing effects, they can be corrected with a design that allows for them. Moreover, they are mitigated by allowing controlled iterations of the price list format. These iterations also increase the precision with which one elicits measures of risk aversion and discount rates for individuals.

Table of Contents

† Centre for Economic and Business Research, Copenhagen, Denmark (Andersen and Lau) and Department of Economics, College of Business Administration, University of Central Florida, USA (Harrison and Rutström). E-mail contacts: SA@CEBR.DK, 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. We are grateful to Melonie Sullivan for comments. All data and instructions are stored in the ExLab Digital Library at http://exlab.bus.ucf.edu.

The literature in experimental economics has been dominated by the use of one important design feature: experimenter-induced values. The explosion of experimental applications in recent decades testifies to the power that controlling via induced values creates.[1] More recently, however, experimental methods have been used to *elicit the homegrown values* of individuals or groups for commodities or projects that exist outside the laboratory. The objective of eliciting values is quite different from the objectives of applications of the induced-values technique. The adjective "homegrown" simply means "not induced," and refers to values that are neither controlled nor known *a priori* by the experimenter. Though more widely applicable, the elicitation of homegrown values is particularly central to the fields of marketing,[2] environmental damage assessment,[3] and the general estimation of individual preferences.[4]

Using laboratory experiments, we examine closely the properties of one procedure which has been widely used to elicit homegrown values: the Multiple Price List (MPL). The MPL is a relatively simple procedure for eliciting values from a subject. In the context of eliciting a willingness to pay for some commodity, it confronts the subject with an array of ordered prices in a table, one per row, and asks the subject to indicate "yes" or "no" for each price. The experimenter then selects one row at random, and the subject's choice is implemented. We examine the behavioral properties of the Multiple Price List (MPL) elicitation institution, as well as some variants on the basic design, in the elicitation of risk attitudes and individual discount rates.

The MPL has several attractions. First, it is easy to explain to subjects. Second, it is easy to implement. Third, it is relatively easy for subjects to see that truthful revelation is in their best interests: if the subject believes that his responses have no effect on which row is chosen, then the

---

[1] Book-length surveys by Davis and Holt [1993], Friedman and Sunder [1994] and Kagel and Roth [1995] predominately report induced-values experiments.

[2] For example, see Hoffman, Menkhaus, Chakravarti, Field and Whipple [1993].

[3] See Shogren [2004] or Harrison [2004] for reviews. In this literature the elicitation of homegrown values for a lab commodity is used either as a proxy for the field environmental good or as a way to calibrate field survey responses.

[4] For example, see Coller and Williams [1999], Harrison, Lau and Williams [2002] and Holt and Laury [2002].

task collapses to a binary choice in which the subject gets what he wants if he answers truthfully.

The MPL design has been employed in three general areas in experimental economics:

- Eliciting risk attitudes: Binswanger [1980][1981] appears to be the first experimental economist to identify risk attitudes using the MPL with real payoffs.[5] It was later used by Murnighan, Roth and Schoumaker [1987][1988], Beck [1994], Gonzalez and Wu [1999], Holt and Laury [2002](HL), Laury and Holt [2002][2004], Eckel and Grossman [2002] and Harrison, Lau, Rutström and Sullivan [2005] (HLRS).
- Eliciting willingness to pay: Kahneman, Knetsch and Thaler [1990] appear to be the first to use it to elicit valuations for a commodity.
- Eliciting individual discount rates: Coller and Williams [1999] appear to have been the first to use the MPL format in this context. Their basic design was employed later by Harrison, Lau and Williams [2002] (HLW), Coller, Harrison and Rutström [2003] and HLRS.

The use of the MPL also has a longer history in the elicitation of hypothetical valuation responses in "contingent valuation" survey settings, as discussed by Mitchell and Carson [1989; p. 100, fn. 14].

The MPL has three possible disadvantages. The first is that it only elicits interval responses, rather than "point" valuations. The second is that subjects can switch back and forth from row to row, implying *potentially* inconsistent preferences. The third is that it could be susceptible to framing effects, as subjects are drawn to the middle of the ordered table irrespective of their true values. We consider each of these disadvantages, propose extensions of the MPL approach which can address each, and evaluate those extensions in controlled laboratory experiments where we elicit measures of risk aversion and discount rates for individuals.

We conclude that although there are framing effects, they can be corrected with a design that allows for them. Moreover, they are mitigated by allowing controlled iterations of the price list format. These iterations also increase the precision with which one elicits measures of risk aversion

[5] The earliest use of the MPL design in the context of elicitation of risk attitudes is, we believe, Miller, Meyer and Lanzetta [1969]. Their design confronted each subject with 5 alternatives that constitute an MPL, although the alternatives were presented individually over 100 trials.

and discount rates for individuals.

In section 1, 2 and 3 we discuss three issues with the use of the MPL: the fact that it generates interval responses rather than point responses, the implications of subjects that "switch" from one row to another more than once, and the possibility of framing effects from the presentation of the entire list of choices. In section 4 we introduce the two valuation tasks considered here: risk attitudes and individual discount rates. Section 5 presents our experimental design, and section 6 m reviews the main results. Section 7 examines a specific puzzle that arose from our findings, and section 8 draws general conclusions.

1. Interval Responses

The problem of interval responses is that one only elicits intervals from the subject rather than point estimates. Thus one does not have as precise a response as might be obtained by some other method that elicits the point response. Since there is some controversy over the ability to elicit valuations too precisely using methods that elicit a point response,[6] it could be that the best one can do anyway is elicit interval responses. For now, we remain agnostic on this issue, although the experiments we undertake can help us address the issue empirically.

There are two methods for addressing the issue of interval response.

The first is simply to use statistical methods that recognize that the response is interval-censored. These methods are an extension of traditional Tobit models, which recognize that a dependant variable may be right or left censored at some fixed value.[7] Tobit models can be extended to allow for right *or* left censoring that varies with the subject. A further extension allows each subject's response to be left-censored *and* right-censored, which is just another way of saying that the subject's response is interval-censored. This statistical approach has been used by Coller and Williams [1999] and the applications of the MPL to elicit discount rates.

[6] See Harrison [1992].
[7] In the original context, expenditures could never be negative.

The second way to address the interval response issue is to extend the MPL to allow more refined elicitation of the true valuation. In order to allow refinements in an efficient manner, one must first ensure that each subject offers a unique and consistent interval choice. The design therefore has to impose uniqueness and consistency, and then allow iterations to refine responses. To see this point, consider the following MPL designs:

- MPL – this is the standard format in which the subject sees a fixed array of paired options and chooses one for each row. It allows subjects to switch back and forth as they like, and has has been used in many experiments already.
- sMPL – Switching MPL varies the standard MPL by asking the subject to simply choose which row he wants to switch at, assuming monotonicity of the underlying preferences to fill out the remaining choices for the subject. This is an important behavioral bridge to the Iterative MPL below, since the latter implicitly assumes such behavior. In all other respects sMPL looks just like the standard MPL.
- iMPL – Iterative MPL extends the Switching MPL to allow the individual to make choices from refined options within the option last chosen. That is, if someone decides at some stage to switch between values of $10 and $20, the next stage of an iMPL would then prompt the subject to refine the values elicited within this interval. When the values being elicited drop below some given perceptive threshold, the program stops iterating.

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

The sMPL is implemented because the iMPL changes the decision from the MPL in two ways: forcing a single switch point in each table, and refining the choice.[8] By comparing MPL and sMPL we can see the pure effect of the first change, and by comparing sMPL and iMPL we can see the pure effect of the second change. We believe that the first implementation of the enforced-single-switching feature of the sMPL was by Gonzalez and Wu [1999]. The theoretical effect of the sMPL format is to impose a strict monotonicity in revealed preferences, as well as enforcing transitivity. These are useful things to check for if one is testing the basic axioms of utility theory, but may be worth imposing if the objective it to elicit consistent responses.[9]

2. Multiple Switch Points

The problem here is that some subjects switch back and forth as they move down the rows of the MPL. This is only a problem for inference if one wants to impose a certain structure on the subject's responses that might not be justified by the underlying theory. For example, few of the existing MPL implementations allow subjects to report indifference. It is quite possible that switching behavior is the result of the subject being indifferent between the options. The implication here is that one simple use a "fatter" interval to represent this subject, defined by the first row that the subject switched at and the last row that the subject switched at. In standard utility theory, this is simply saying that preferences are only required to be weakly convex rather than strictly convex.

However, this interpretation of possible switching behavior in the MPL institution implies that we should allow an explicit indifference option in the implementations of the sMPL and iMPL institutions. Comparison of behavior from comparable implementations of the MPL and sMPL

[8] In a literal sense one could design an iMPL institution that did not built in the sMPL features at each stage. The problem is that it could be inefficient and confusing to subjects. One would have to define the second stage over the largest set of switch points in the prior stage, or else undertake second stages separately for each and every switch point. One of the great attractions of the MPL procedure in the first place is the relative transparency of the task to subjects, and we are loathe to burden that transparency any more than is needed.

[9] There is a parallel in the field of elicitation of values using incentive-compatible procedures, such as a Vickrey auction. One can test if subjects understand the incentive compatibility property, or one can explain it. The latter is appropriate if the goal is simply to elicit true valuations and one is willing to maintain the assumption that subjects follow the logic of telling the truth. In short, one does not have to test everything in every experimental design.

institution (and the *first* stage of the iMPL institution) will allow an evaluation of the behavioral effect of constraining preferences to be explicit about indifference.

3. Framing Effects

A natural concern with the MPL is that it might encourage subjects to pick a response in the middle of the table, independent of true valuations. There could be a psychological bias towards the middle, although that is not obvious *a priori*. More to the point in a valuation setting, the use of specific values at either end of the table could signal to the subject that the experimenter believes that these are reasonable upper and lower bounds. In some tasks, such as the risk elicitation task of HL, the values are bounded by the laws of probability between 0 and 1, so this is less likely to be a factor compared to the pure psychological anchor of the middle row.

One solution to this task which we find unattractive is to randomize the order of the rows.[10] This is unattractive for two reasons. First, if there is a purely psychological anchoring effect towards the middle, this will do nothing but add noise to the responses. Second, the valuation task is fundamentally harder from a cognitive perspective if one shuffles the order of valuations across rows. Such a task may be of interest in some field settings, although we cannot imagine any. We are not interested in testing if subjects have the ability to re-order the valuations *and* identify their preferred valuation. Our interest is only in the latter question.

Framing effects can be relatively easily evaluated by varying the cardinal scale of the basic MPL table, or by varying the number of intervals within a given cardinal range. For example, assume one were eliciting individual discount rates and the initial cardinal scale was between 1% and 50%. Give this MPL task to one set of subjects, and then give an MPL task in which there were additional rows going up to 100%. If there is a difference in response between the two samples, it will be easy to identify statistically and then to correct for it in the data analysis.

We would not be surprised to find framing effects of this kind. They do not necessarily

[10] For example, see Kirby and Maraković [1996] and Kirby, Petry and Bickel [1999].

indicate a failure of the traditional economic model, so much as a need to recognize that subjects in a lab setting use all available information to identify a good valuation for a commodity.[11] Thus it is critical to be able to estimate the quantitative effect of certain frames and then correct for them in subsequent statistical analysis. In other words, the existence of a bias due to a particular frame should ideally just lead to a statistical "calibration" of the responses to correct for the particular frame. We devise a test for framing effects by varying the cardinal scale of the MPL in both the risk aversion task and in the discount rate task.

Two asymmetric frames are developed: In the risk aversion task, the *skewHI* treatment offers initial probabilities of (0.3, 0.5, 0.7, 0.8, 0.9 and 1), while *skewLO* offers initial probabilities of (0.1, 0.2, 0.3, 0.5, 0.7, and 1). This treatment yields 6 decision rows in Level 1 of the iMPL, as opposed to the 10 rows in the symmetric frame.[12] In the discount rate task, the *skewHI* treatment offers initial annual interest rates of (15%, 25%, 35%, 40%, 45%, and 50%), while the *skewLO* treatment offers annual interest rates of (5%, 10%, 15%, 25%, 35%, and 50%). The symmetric treatment offers 10 rows with annual interest rates between 5% and 50%.

4. Specific Valuation Tasks

A. Risk Aversion

HL devise a simple experimental measure for risk aversion using a multiple price list (MPL) design. Each subject is presented with a choice between two lotteries, which we can call A or B. Table 1 illustrates the basic payoff matrix presented to subjects. The first row shows that lottery A offered a 10% chance of receiving $2 and a 90% chance of receiving $1.60. The expected value of

[11] See Harrison, Harstad and Rutström [2004] for further discussion of these channels.

[12] The skewed frames will affect the implementation of the iMPL. In the symmetric frame, all intervals in Level 1 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 in the first level vary in size, a third level is required to bring choices down to this level of precision, and the number of decision rows in Level 3 depends on the width of the interval in Level 1 at which the subject switches. We use the same procedure in the discount rate task, and the threshold at the minimal bi-section interval is 0.1 af a percentage point.

this lottery, EV^A, is shown in the third-last column as \$1.64, although the EV columns were not presented to subjects.[13] Similarly, lottery B in the first row 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. 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. Arguably, the last row is a test that the non-satiated 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.[14]

In our design we deliberately follow the field experiments of HLRS and ask each subject to respond to four risk elicitation tasks. The prizes in each task differ, so that there is a finer grid of risk attitudes revealed.[15] We depart from HLRS in one respect: we randomize the order of the four tasks across subjects, and then randomly decide which one to play out.[16] This will allow us to check for a possible order effect in the risk attitudes elicited in the field by HLRS; there is evidence from Harrison, Johnson, McInness and Rutström [2003a] that the experiments of HL do exhibit order

[13] There is an interesting question as to whether they should be provided. Arguably the subjects are trying to calculate them anyway, so providing them avoids a test of the joint hypothesis that "the subjects can calculate EV in their heads and will not accept a fair actuarial bet." On the other hand, providing them may cue the subjects to adopt risk-neutral choices. The effect of providing EV information deserves empirical study.

[14] Following Rabin [2000], there are some specifications of expected utility theory for which a finding of risk aversion at these levels of income implies incoherent behavior. This implication does not apply if expected utility theory is defined over income earned during the experiment, rather than over terminal wealth (Cox and Sadiraj [2004]).

[15] The four sets of prizes are as follows, with the two prizes for lottery A listed first and the two prizes for lottery B listed next: (A1: 2000 DKK, 1600 DKK; B1: 3850 DKK, 100 DKK), (A2: 2250 DKK, 1500 DKK; B2: 4000 DKK, 500 DKK), (A3: 2000 DKK, 1750 DKK; B3: 4000 DKK, 150 DKK), and (A4: 2500 DKK, 1000 DKK; B4: 4500 DKK, 50 DKK). At the time of the experiments, the exchange rate was approximately 6.7 DKK per U.S. dollar, so the prizes range from approximately \$7.5 to \$672.

[16] The large incentives and budget constraints precluded paying all subjects, so each subject is given a 10 percent chance to actually receive the payment associated with his decision.

effects that are statistically and substantively significant. In our statistical analysis we control for "task effects" by adding binary indicator variables. These will capture learning effects as well as order effects.[17]

B. Individual Discount Rates

The basic question used to elicit individual discount rates is extremely simple: do you prefer \$100 today or \$100+*x* tomorrow, where *x* is some positive amount? If the subject prefers the \$100 today then we can infer that the discount rate is higher than *x*% per day; otherwise, we can infer that it is *x*% per day or less. The format of the field experiments in HLRS and HLW modified and extended this basic question in five ways, which we follow.

First, a number of such questions were posed to each individual, each question varying *x* by some amount. When *x* is zero we would obviously expect the individual to reject the option of waiting for no rate of return. As we increase *x* we would expect more individuals to take the future income option. For any given individual, the point at which they switch from choosing the current income option to taking the future income option provides a bound on their discount rate. That is, if an individual takes the current income option for all *x* from 0 to 10, then takes the future income option for all *x* from 11 up to 100, we can infer that their discount rate lies between 10% and 11% for this time interval. The finer the increments in *x*, the finer will we be able to pinpoint the discount rate of the individual.

Second, HLW used an MPL format and HLRS used an iMPL format, simultaneously posing several questions with varying values of *x*, and selecting one question at random for actual payment after all responses have been completed by the individual. In this way the results from one question do not generate income effects which might influence the answers to other questions.

[17] In the context of the experimental design of HL the order of the task was confounded with a scaling of all lottery prizes, as well as changes from hypothetical to real payoffs. Thus it is critical in that design, if one wants to make independent inferences about the effects of scaling prizes or the use of hypothetical rewards, to identify any separate effect from order effects. In our design there are no such treatments confounded with order.

Third, the field experiments provided choices between two future income options rather than one "instant income" option and one future income option. For example, we offer \$100 in one month and \$100+x in 7 months, interpreting the revealed discount rate as applying to a time horizon of 6 months. This avoids the potential problem of the subject facing extra transactions costs[18] with the future income option. If the delayed option were to involve greater transactions costs, then the revealed discount rate would include these subjective transactions costs. By having both options entail future income we hold these transactions costs constant.

Fourth, the field experiments provided respondents with the interest rates associated with the delayed payment option. This is an important control feature if field investments are priced in terms of interest rates. If subjects are attempting to compare the lab investment to their field options, this feature may serve to reduce comparison errors since now both lab and field options are priced in the same metric.

Subjects in the 6-month horizon treatment were given payoff tables as illustrated in Table 2. They were told that they must choose between payment Options A and B for each of the 10 payoff alternatives. Option A was 3000 DKK in all sessions, payable in 1 month. Option B paid 3000 DKK + x DKK in 7 months, where x ranged from annual rates of return of 5% to 50% on the principal of 3000 DKK, compounded quarterly to be consistent with general Danish banking practices on overdraft accounts. The payoff tables provided the annual and annual effective interest rates for each payment option and the experimental instructions defined these terms by way of example.

In our design we deliberately follow the field experiments of HLRS and ask each subject to respond to three of the six horizons employed by HLRS: the one month, four month and six month horizons.[19] We depart from HLRS in one respect: we again randomize the order of the three tasks across subjects. This will allow us to check for a possible pure order effect in the discount rates elicited in the field by HLRS. The exchange rate at the time of the experiments in October 2003 was

[18] Including the possibility of default by the experimenter.

[19] We randomly decide which task to play out, and each subject is given a 10 percent chance to actually receive the payment associated with his decision.

approximately 6.7 DKK per US dollar, so the base amount of option A converts to approximately $450.

5. Our Experiments

Our basic design explores each of the components of the MPL elicitation format reviewed above. We examine the performance of the three MPL institutions (MPL, sMPL and iMPL) for each of the three framing conditions ("skew low", symmetric, and "skew high") and the two types of valuation tasks (risk aversion and discount rates). Thus we have a 3×3×2 design. The first two treatments are implemented between-subjects, so that any one subject only experiences one of the MPL institutions and one of the frames.[20] The last treatment is implemented within-subjects, such that each subject faces 4 risk aversion tasks with varying stakes and 3 discount rate tasks with varying horizons.

In addition, we have three treatments that are applied equally to all subjects in each session. One is a randomization of their initial endowment. Each subject received a guaranteed 250 DKK to participate. In addition, we randomly assign them an extra amount between 10 and 100 DKK, chosen from a discrete uniform distribution in increments of 10 DKK. Following Rutström [1998], the purpose of this treatment is to determine if there are endowment or "house money" effects on behavior, at least within the range considered here. No subject knows the additional amount received by any other subject, but will know that the same random process was applied to all subjects. The second treatment is a randomization of the four risk aversion tasks. There are 24

[20] To avoid session effects interacting with treatment effects, we would have to provide instructions on the three elicitation formats in one of several ways. Either they would have to be completely private, using fully-computerized or written instructions. Thus we could randomly assign subjects to treatment within each session. Or we could have designed our instructions so that they would introduce subjects to all three institutions, and then just implement one at random with a given subject. Neither of these alternatives were attractive, since we wanted to use public instruction to ensure that subjects were paying attention rather than relying solely on their reading comprehension and some quiz questions. Introducing all three formats could have introduced a treatment effect for MPL and sMPL itself (but not for iMPL, since the iMPL logic virtually requires that one explain the MPL and sMPL logic). We decided to assume that any session effects are picked up by the mix of observable characteristics identifying the sample that we have in each session, and controlling for those characteristics when comparing treatments. This is a good "second best" to controlling for session effects, but a reasonable one given the alternatives.

different sequences of risk aversion tasks, but only 4 sequences are chosen. Each subject is assigned to one of the four sequences randomly at each session.[21] The last treatment is a randomization of the three IDR tasks. There are six sequences of horizons in the IDR tasks. All six sequences are used in the experiment, and are randomly assigned to each subject.

We recruited 100 subjects from the University of Copenhagen and the Copenhagen Business School in October 2003, spread across 9 sessions. All subjects were recruited using the ExLab software.[22] The sessions were announced in 7 different lectures. At each lecture an announcement of the experiment was read aloud, and subjects were asked to enrol for the experiment by accessing ExLab through the Danish web page for this project.

The experiments were conducted in October of 2003. Of the 100 subjects recruited, 90 showed up for the experiment evenly spread across the 9 sessions. Although several non-students participated, 74 out of the 90 subjects were students. Ages varied from 18 to 32 years, averaging 22.7 years, and 27% were female. The sessions where conducted in the same manner as HLRS. Because the sessions lasted for two hours, light refreshments were provided before the start of the session.

6. Results

Our analysis examines the effect of our treatments on the average measures of risk aversion and individual discount rates elicited. Since we have information on several demographic characteristics we can also investigate these potential correlates.

[21] The four different sequences of risk aversion tasks are (1, 2, 3, 4), (2, 4, 1, 3), (3, 1, 4, 2) and (4, 3, 2, 1).

[22] This recruitment software is available for academic use at http://exlab.bus.ucf.edu. In addition, all instructions are provided for review at the ExLab Digital Library at the same location.

A. Average Measures of Risk Attitudes

Figure 1 shows the observed distribution of risk attitudes in the lab experiment, using both raw mid-point of the elicited intervals and estimated values. All mid-points and estimated values are generated with the final iteration of MPL formats; the statistical model used for the estimated values is explained below. In the MPL and sMPL the initial and final iteration coincide, of course, but there may be several levels of iterations in the iMPL. For comparability, these distributions only reflect the symmetric menu treatment. Using CRRA as the characterization of risk attitudes,[23] a value of 0 denotes risk neutrality, negative values indicate risk-loving, and positive values indicate risk aversion. Thus we see risk aversion in the lab: the mean CRRA coefficient is 0.72 and the median is 0.70. Using raw midpoints, only 12 of 90 subjects appear to exhibit any risk-preference in any task, and none exhibit risk-loving preferences for all tasks.

Table 3 displays estimates from a panel interval regression model of the elicited CRRA values.[24] This is the statistical model used to generate the estimated CRRA responses shown in the bottom panel of Figure 1.[25] The coefficients in the regression model can be interpreted as the marginal effect from each variable compared to the default, which in this case is a male, married, does not own home or apartment, non-student, unskilled, middle class, and lives in city of less than 20.000 people.

We find that responses from the iMPL treatments are associated with higher risk aversion. This is significant for both initial and final answers, with iMPL increasing risk aversion by 0.36 (p-value = 0.001) for final answers. We find no effect from sMPL on comparable choices at the first stage, implying that enforcing a switching point[26] had no systematic effect. Figure 2 shows the

[23] The specific functional form used is $U(m) = (m^{1-r})/(1-r)$, where r is the CRRA coefficient. When r =1, $U(m) = \ln(m)$.

[24] Four subjects chose either indifference or option A in the last row of the MPL, so no CRRA bounds can be calculated for them. One subject chose option A for all lotteries for two of the tasks. Thus we report 354 observations in the regression, instead of the 360 observations we would have if all subjects had responses that could have been used.

[25] Each prediction is conditional on the actual characteristics of the subjects, and the estimated coefficients from the model reported in Table 3.

[26] And hence enforcing strict monotonicity and transitivity.

estimated CRRA values of subjects for each MPL format. We observe in the bottom panel the effect of iMPL format: an increase in elicited risk aversion.

The framing of the decision table does have a significant effect on responses, the effect from *skewLO* to lower elicited CRRA by 0.23 with a *p*-value of 0.05. Figure 3 shows the predicted responses for each frame, and the slight effect from the framing designed to lower elicited CRRA.

There is an order effect on the CRRA coefficient from the variation of sequence of lottery prizes across the four tasks. There is a significant difference between the reference Task 1 and the last task, Task 4. This is significant at the 0.3% level, with subjects having a higher CRRA of 0.14 on average. This differs from the findings in HLRS, who find order effects from the last three tasks. In our experiment we randomize the order of tasks, and are thus able to test for pure order effects.[27] Figure 4 shows the predicted values of CRRA under each task, and the effect of the last task in the bottom panel.

Since there are variations in responses across subjects, it is important to test if these response variations are captured by observable characteristics such as demographics. Only "single" status is significant, and then only at the 5.5% level: the effect is for single subjects to have a lower risk aversion by about 0.24 on average. In general, we see far less heterogeneity in our sample than the comparable findings in HLRS, no doubt reflecting the greater homogeneity of our sample.

We thus find that the iMPL format is associated with an increase in the elicited CRRA coefficient, that there is an effect from framing the task to generate a lower level of risk aversion, and that there is an order effect from the last task.[28]

---

[27] It is also possible that there is some effect from the specific prizes included in each lottery choice task. This could occur if CRRA is not an appropriate characterization of risk attitudes over the domain of prizes, and subjects exhibit non-constant RRA for some of the prize levels. Evidence for variations in RRA with scale changes in payoffs was one of the primary conclusions of Holt and Laury [2002], and is also supported by the statistical analysis of Harrison, Lau and Rutström [2004]. To test for an effect here, we also included dummy variables for each of the specific prize sets in the field experiments of HLRS. These had no effect at all on the estimates of order effects we report here. But we do detect a statistically significant effect of the prize structure in two of the four lotteries, consistent with the non-constant RRA story. Specifically, RRA is 0.11 higher for the second field lottery pair, relative to the estimates for the first field lottery pair.

[28] We also check for interaction effects of the main treatments, pooling MPL and sMPL responses since they lead to identical choices. First we analyze whether there is an interaction effect between formats and framing, and can reject the joint hypothesis of no interaction at the 9% level. Testing each interaction

B. Average Measures of Individual Discount Rates

Figure 5 displays the elicited discount rates for our subjects in the lab, using the mid-point of the final interval selected as well as estimated values.[29] The discount rates are pooled across all horizons, and these distributions only reflect the symmetric menu treatment. We observe variations of elicited discount rates across subjects, with a mean of 29.2%, a median of 27.9%, and a standard deviation of 15.4%.[30]

Table 4 reports the results from a panel interval regression of the elicited discount rates, controlling for horizon, multiple price list formats, framing effects, order effects and individual demographics. This model uses panel data since each subject provided three interval responses, one for each horizon. The regression shows that the 4 and 6 months horizons have significantly lower discount rates than the reference horizon, which is 1 month. These rates are between 7 and 9 percentage points lower than the 1 month rates. The elicited discount rates do not vary across the two longer horizons.[31]

We find no effect of iMPL and sMPL formats on either initial or final responses.[32] Figure 6 shows the predicted IDR responses by subjects for each format. We also find that the framing of the table initially presented to subjects does not have a significant effect on responses.[33] Figure 7 shows

---

separately, we find that there may be an interaction between formats and high framing, but that there is none for low framing. Since the interaction between iMPL and high framing is potentially confounded with a (single) session, this could be a session effect rather than an interaction effect. Our priors are that this is not a session effect, given the similarities in the recruitment and conduct of all sessions. Looking at interaction effects between formats and task order, we find an interaction effect of iMPL on the final task. Using a Wald test, we can reject the joint hypothesis of no interaction only at a 10% level between formats and order of tasks. We identify an effect from iMPL on Task 4 that is significant and associated with an increased CRRA of 0.29 on average. We believe that this is either a learning effect or a fatigue effect of the iMPL tables. In Section 7 we report evidence from additional experiments that supports the hypothesized learning effect.

[29] We use elicited responses in the laboratory which are assumed to be uncensored by market prices. This assumption is consistent with the view that subjects do not integrate income earned in the laboratory with personal financial wealth. See Cox and Sadiraj [2004] for a discussion of the asset integration hypothesis.

[30] These values are virtually identical to those found in the field by HLRS: they estimated a mean of 24.2%, a median of 24.5%, and a standard deviation of 15.7%.

[31] This conclusion follows from a Wald test with a significance level of 16%, implying that we cannot reject the null hypothesis that the two coefficients in question are the same value.

[32] This conclusion is confirmed by a test of the joint hypothesis that the effect from iMPL and sMPL is zero. A Wald test does not reject the null hypothesis with a *p*-value of 0.48.

[33] We cannot reject the joint hypothesis that the joint effect from both asymmetric framings is zero. A Wald test does not reject this with a *p*-value of 0.66.

the predicted IDR responses by subjects for each framing.

We observe an order effect on the elicited discount rate from the variation of sequence of horizons. There is a significant difference between the discount rates elicited in reference Task 1 and the subsequent tasks. Task 2 is associated with an average discount rate that is higher by 3.5 percentage points, which is significant at the 3.5% level. In our experiment we randomize the order of tasks, and these effects are thus pure order effects. We cannot reject the hypothesis that the average rates for Task 2 and Task 3 are the same (*p*-value = 0.42). This suggests that there was some learning effect but only after the first task. Figure 8 shows the predicted IDR responses by subjects for each task.

Since there are variations in responses across subjects, it is important to test if these response variations are captured by observable characteristics such as demographics. None of the socio-demographic variables are significant at conventional levels. However, we do find an experimenter effect, although this effect is significant only at the 6.8% level.

We thus find that the iMPL format has no discernible effect on the elicited discount rates, that there is no effect from framing on discount rates, and that there is an order effect on the second task after the initial task.[34]

7. Does the iMPL Format Increase Elicited Risk Aversion?

Our results raise one serious puzzle: the apparent effect of the iMPL format on risk attitudes. The statistical results presented above suggest that it has a statistically significant effect on risk attitudes, increasing them by approximately 0.36 in terms of a CRRA characterization and

[34] We also check for interaction effects of the main treatments, pooling MPL and sMPL responses since they lead to identical choices. First we analyze whether there is an interaction effect between formats and framing, and cannot reject the joint hypothesis of no interaction (a Wald test has a *p*-value of 0.60 for the joint hypothesis of no interactions between format and framing). Looking at interaction effects between formats and task order, we find an interaction effect of the *non*-iMPL formats in the last two tasks. Using a Wald test, we cannot reject the joint hypothesis of no interaction effect between *all* formats and order of tasks. But we do find an effect from the *non*-iMPL responses on Task 2 and Task 3 that is significant at the 2.2% and 12.8% level, respectively. The non-iMPL responses are associated with an increase in the average elicited discount rate of 4.7 and 3.1 percentage points for Task 2 and Task 3, respectively.

relative to the values generated by the MPL and sMPL formats (see Table 3). What might be driving this apparent effect?

First, this could be due to some outliers. Figure 9 displays individual responses for virtually all of the subjects in our experiments. We show the mid-point of the elicited CRRA interval, at the end of any iterations in the iMPL format. We also drop a handful of subjects that exhibited risk-loving behavior, since that distorts the pictures needlessly. The iMPL sessions are on a diagonal, in panels 3, 5 and 7. Three subjects stand out to the eyeballs as being outliers: subject 949 in the symmetric iMPL session (panel 5), and subjects 967 and 977 in the *SkewLO* iMPL session (panel 7). If one includes dummy variables for each of these three, and repeats the statistical analysis underlying Table 3, the estimated impact of the iMPL is reduced from 0.36 to 0.21. In addition, the *p*-value on the null hypothesis test that this coefficient is equal to zero increases from 0.001 to 0.032, and the 95% confidence interval spans an effect of 0.02 to +0.39. Thus, these three individual subjects do play some role in explaining the significance of the effect. However, such *post hoc* specifications can be dangerous, since one can potentially search over individual subject responses to generate any preferred inference.[35]

Second, this effect could be due to there being some effect of the iMPL format on residual variance of the elicited CRRA measure that is being picked up as a change in the average CRRA. Inspection of Figure 9 lends some support to this possibility, since the distributions for the iMPL cells seem to have a higher variance. We can check this hypothesis by including a multiplicative heteroskedasticity term in an interval regression model, and allowing the variables for the elicitation format, the skewness frame and the task order to have an effect.[36] We find that although the iMPL format does significantly increase residual variance, this does not change the conclusion with respect to the effect of the iMPL frame on mean CRRA.

[35] On the other hand, they do informally point to the value of larger samples in the iMPL cells.

[36] This heteroskedasticity extension cannot be estimated with the random effects specification to account for possible unobserved individual effects. So we pool responses over all subjects and tasks. We do allow for "robust" standard errors using the Huber-White correction, extended to allow for possible clustering on the responses of the same subject.

Third, the fact that there was no comparable effect in the discount rate experiments suggests that there might be some learning effect present. That is, the subjects completed the risk aversion tasks prior to the discount rate tasks, and the responses in the risk aversion tasks might have been confounded by some subjects still learning about the mechanics or properties of the iMPL format. By the time those subjects reached the discount rate tasks, they had presumably learned what they were going to learn, and there was no observed effect. This hypothesis is particularly interesting since the task itself involves evaluating risk attitudes, so if the subject is uncertain about the procedures as well as the underlying lotteries, there is a compound lottery being evaluated. Hence, under reasonable conditions on the effect of "background risk," one would expect to see higher measures of risk aversion from those subjects (e.g., Gollier [2001; chs. 8,9]).

To test this hypothesis we conducted a further series of experiments in the United States where we varied the sequence in which the risk aversion task appeared. In some cases it was the first task, and in others it was the second task.[37] In each case the same four tasks were employed, just with different orders (within the risk aversion tasks) and sequences in relation to other valuation tasks. All other basic procedures remained the same, and the same computerized interface was employed in all experiments (with instructions provided in English instead of Danish). There were several additional procedural changes, so we are careful not to pool data from the U.S. and Danish experiments without appropriate controls.[38] In February of 2004 we recruited 116 subjects from

[37] The first task was a series of valuation exercises in which we elicited willingness to pay for four commodities. As an aside, these experiments also allow additional insight into the interpretation of the task order controls. Since we varied task order and prizes, as well as whether these responses were elicited after the subject was as familiar with the iMPL as they were likely to be, we can see if familiarity reduced the effect of task order. It did, so we can conclude that the task order effect was probably due to initial learning effects with the iMPL format.

[38] At the risk of being very literal, there were five differences. First, in the examples provided to subjects. In Denmark we used bingo balls and publicly went through the examples. In the U.S. we used dice, and subjects were introduced to the dice individually, but in every other aspect the examples were the same. Second, with the candy trainer. In Denmark we used bingo balls in public. We then rolled the 10-sided die to see if subjects received the candy. If a 0 was not rolled, we would not show them their choices. In the U.S. we rolled the 10-sided dice first to see whether they won. Whether they won or not, we went through their choices. We made them explain what they could have won conditional on their choices and the roll of the dice. This change was made to ensure that every subject saw what their choices would have been, and not just those picked by chance as in the Danish procedures. Third, in the tasks themselves, we followed the procedures just mentioned for the trainer. We also selected the row and the outcome individually, rather than

large lectures in the College of Business Administration at the University of Central Florida. Of these 116, we ran the risk aversion task first for 73 subjects. Figure 10 displays the comparable distribution of elicited CRRA values from the Danish and U.S. experiments. Average CRRA is estimated to be 0.70, with a standard deviation of 0.56 and median of 0.66. The 95% confidence interval spans a low of -0.169, consistent with some slight risk-loving behavior, and a high of 1.65. Thus subjects were slightly less risk averse in the U.S., but with a wider range.

Using the same statistical analyses as we used for the Danish data, we find that there is a small effect from the iMPL format when risk aversion task comes second, but that it is not statistically significant. Table 5 reports the detailed results of the main statistical analysis of the U.S. responses. We find that average CRRA is 0.19 higher when the risk aversion task is second, with a *p*-value of 0.14 and a 95% confidence interval spanning -0.06 and +0.44. In these experiments there is still an effect of the iMPL to increase elicited risk aversion by 0.08 on average, but it is no longer statistically significant. The 95% confidence spans -0.19 and +0.35, so there does not appear to be an effect from iMPL in the U.S. experiments. This conclusion also holds if we restrict the analysis to the 73 subjects who responded to the risk aversion task first, as in the Danish experiments.

Fourth, the result from the Danish experiments could just be due to small samples, allowing a few outliers to distort inferences. We can check this hypothesis by pooling the data from the U.S. and Danish lab experiments, and controlling for possible differences between the two.[39] If we do

in a public manner. Fourth, we employed an in-sample design for the skewness frame in the U.S., whereas in Denmark each subject saw only one such frame. Fifth, the prizes used in the U.S. were downscaled by 13.5 to 1 in each task, using the monthly average exchange rate at the time of the experiments. Since subjects in the U.S. had a chance of winning in each of the four tasks, the overall average payment for risk tasks was effectively only a downscale of 3.3 to 1. Thus, in practice, the average payment to the Danish subjects in this task was 220 DKK, and in the U.S. would have been 220 DKK ÷ 3.3 = 66 DKK, or roughly USD 11. We reduced payments in the U.S. for overall budgetary reasons. Finally, in terms of the order of the tasks, in the U.S. design we used willingness to pay tasks before the risk aversion task for some sessions, whereas in Denmark the risk aversion tasks were always first.

[39] The dangers of pooling here are the same as the dangers of making cross-country comparisons in general: see Botelho, Harrison, Hirsch and Rutström [2005] for extensive discussion of the econometric issues in the context of bargaining experiments. Apart from controlling for all standard observable characteristics of the subject, such as age and sex, we include country dummies and interactions between the iMPL treatment and country since that is the focus of the analysis. Neither is statistically significant, or changes our conclusions with respect to the effect if iMPL on behavior. We also control the possible effects of country and iMPL procedure on residual variance, and find that it does not change our conclusion with respect to the

this the effect of the iMPL format is reduced by more than 50%, and is statistically insignificant. Just pooling with the U.S. data in which the risk aversion task came first, since it is most directly comparable to the Danish data, we estimate that the iMPL format increases average CRRA by only 0.10 on average, and that the effect is not statistically significant. The *p*-value on this estimated effect is 0.37, and the 95% confidence intervals spans -0.12 to +0.31. The effect of iMPL on elicited risk aversion is even less significant if we pool in all of the U.S. data, adding controls for the order of the risk aversion tasks.

8. Conclusions

Behavioral responses to the MPL and iMPL formats are sensitive to some of the concerns raised about them, but these effects are amenable to experimental control and statistical correction.

Framing can influence elicited values, but in ways that are relatively easy to correct for statistically if allowed for in the design. We expect that such effects will be allowed for by design, much as one routinely checks for and corrects for possible order effects.

There appear to be some order effects from the sequencing of specific risk or discount rate elicitation tasks.[40] These are relatively small in quantitative size in the case of discount rates, even though they are statistically significant. The evidence from auxiliary experiments conducted to examine this issue suggests that they are due to learning effects, and disappear as familiarity with the iMPL format is accumulated.

The precision of our estimates of risk aversion and discount rates is increased with the iMPL format if one uses the direct estimate obtained from the responses of the subject. This improvement in precision is by design, of course, and can be extremely valuable when using these elicited values as

impact of the iMPL procedure. We do not conclude that one can pool the Danish and U.S. data for all purposes, but that any differences in the two do not influence the inferences we make about the iMPL procedure.

[40] These order effects are different from the possibility of sequencing effects from undertaking all of the risk tasks prior to all of the discount rate elicitation tasks. We do examine the effects of sequencing in our U.S. experiments, discussed in section 7.

controls in paired, within-subjects experiments. It allows more precise tests of hypotheses that are conditioned on individual risk aversion or discount rates.

Table 1: Payoff Matrix from the Holt and Laury Risk Aversion Experiments

Lottery A				Lottery B						
p($2)		p($1.60)		p($3.85)		p($0.10)		EV^A	EV^B	Difference
0.1	$2	0.9	$1.60	0.1	$3.85	0.9	$0.10	$1.64	$0.48	$1.17
0.2	$2	0.8	$1.60	0.2	$3.85	0.8	$0.10	$1.68	$0.85	$0.83
0.3	$2	0.7	$1.60	0.3	$3.85	0.7	$0.10	$1.72	$1.23	$0.49
0.4	$2	0.6	$1.60	0.4	$3.85	0.6	$0.10	$1.76	$1.60	$0.16
0.5	$2	0.5	$1.60	0.5	$3.85	0.5	$0.10	$1.80	$1.98	-$0.17
0.6	$2	0.4	$1.60	0.6	$3.85	0.4	$0.10	$1.84	$2.35	-$0.51
0.7	$2	0.3	$1.60	0.7	$3.85	0.3	$0.10	$1.88	$2.73	-$0.84
0.8	$2	0.2	$1.60	0.8	$3.85	0.2	$0.10	$1.92	$3.10	-$1.18
0.9	$2	0.1	$1.60	0.9	$3.85	0.1	$0.10	$1.96	$3.48	-$1.52
1	$2	0	$1.60	1	$3.85	0	$0.10	$2.00	$3.85	-$1.85

Note: The last three columns in this table, showing the expected values of the lotteries, were not shown to subjects.

Table 2: Payoff Table for 6 Month Time Horizon

Payoff Alternative	**Payment Option A** (pays amount below in 1 month)	**Payment Option B** (pays amount below in 7 months)	**Annual Interest Rate** (AR, in percent)	**Annual Effective Interest Rate** (AER, in percent)	**Preferred Payment Option** (Circle A or B)
1	3,000 DKK	3,075 DKK	5	5.09	A B
2	3,000 DKK	3,152 DKK	10	10.38	A B
3	3,000 DKK	3,229 DKK	15	15.87	A B
4	3,000 DKK	3,308 DKK	20	21.55	A B
5	3,000 DKK	3,387 DKK	25	27.44	A B
6	3,000 DKK	3,467 DKK	30	33.55	A B
7	3,000 DKK	3,548 DKK	35	39.87	A B
8	3,000 DKK	3,630 DKK	40	46.41	A B
9	3,000 DKK	3,713 DKK	45	53.18	A B
10	3,000 DKK	3,797 DKK	50	60.18	A B

Table 3: Statistical Model of Risk Aversion Responses

Random-effects interval regression,
with the final CRRA interval chosen by the subject as the dependent variable.

N=354, based on 90 subjects.

Variable	Description	Estimate	Standard Error	*p*-value	Lower 95% Confidence Interval	Upper 95% Confidence Interval
Constant		0.84	0.38	0.03	0.10	1.58
impl	iMPL format	0.36	0.11	0.00	0.15	0.58
smpl	sMPL format	0.02	0.11	0.88	-0.19	0.23
skewLO	SkewLO frame	-0.23	0.12	0.05	-0.46	-0.00
skewHI	SkewHI frame	-0.01	0.11	0.96	-0.22	0.21
endowment	Random initial endowment	-0.00	0.00	0.34	-0.00	0.00
Task2	Second task	0.04	0.05	0.46	-0.06	0.13
Task3	Third task	0.02	0.05	0.74	-0.08	0.11
Task4	Fourth task	0.14	0.05	0.00	0.05	0.24
experimenter	Experimenter effect	-0.09	0.10	0.34	-0.28	0.10
female	Female	0.09	0.10	0.39	-0.12	0.30
single	Lives alone	-0.24	0.13	0.05	-0.49	0.00
nhhd	Number in household	0.00	0.08	0.97	-0.15	0.15
owner	Owns home or apartment	0.11	0.15	0.46	-0.18	0.40
student	Student	-0.04	0.12	0.72	-0.28	0.19
skilled	Some post-secondary education	-0.04	0.15	0.80	-0.32	0.25
longedu	Substantial higher education	-0.03	0.13	0.82	-0.28	0.22
IncLow	Lower level income	0.05	0.17	0.76	-0.29	0.39
IncHigh	Higher level income	-0.08	0.22	0.73	-0.50	0.35
copen	Lives in Copenhagen area	0.07	0.23	0.76	-0.38	0.52
city	Lives in larger city of 20,000 or more	0.01	0.28	0.98	-0.54	0.55
σ_u	Standard deviation of individual effect	0.36	0.03	0.00	0.30	0.43
σ_e	Standard deviation of residual	0.29	0.01	0.00	0.26	0.31

Notes: Log-likelihood value is -561.2; Wald test for null hypothesis that all coefficients are zero has a χ^2 value of 44.41 with 20 degrees of freedom, implying a *p*-value of 0.0013; fraction of the total error variance due to random individual effects is estimated to be 0.617, with a standard error of 0.051.

Legend: Most variables have self-evident definitions. 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 4: Statistical Model of Individual Discount Rate Responses

Random-effects interval regression,
with the final discount rate interval chosen by the subject as the dependent variable.

N=270, based on 90 subjects.

Variable	Description	Estimate	Standard Error	*p*-value	Lower 95% Confidence Interval	Upper 95% Confidence Interval
Constant		51.55	16.11	0	19.97	83.14
horizon4	4 months horizon	-6.74	1.67	0	-10.01	-3.48
horizon6	6 months horizon	-9.07	1.66	0	-12.32	-5.82
impl	iMPL format	1.85	4.64	0.69	-7.24	10.95
smpl	sMPL format	-3.85	4.52	0.39	-12.71	5.01
skewLO	SkewLO frame	3.54	4.86	0.47	-5.98	13.06
skewHI	SkewHI frame	3.85	4.59	0.4	-5.14	12.84
endowment	Random initial endowment	-0.01	0.07	0.9	-0.14	0.12
Task2	Second task	3.51	1.66	0.03	0.25	6.77
Task3	Third task	2.17	1.65	0.19	-1.06	5.4
experimenter	Experimenter effect	-7.37	4.03	0.07	-15.27	0.54
female	Female	2.45	4.42	0.58	-6.21	11.12
single	Lives alone	-2.5	5.35	0.64	-12.98	7.99
nhhd	Number in household	3.52	3.21	0.27	-2.78	9.82
owner	Owns home or apartment	-7.83	6.25	0.21	-20.08	4.42
student	Student	-6.3	4.98	0.21	-16.06	3.46
skilled	Some post-secondary education	-6.86	6.19	0.27	-18.98	5.27
longedu	Substantial higher education	-8.09	5.38	0.13	-18.64	2.47
IncLow	Lower level income	-3.57	7.38	0.63	-18.04	10.9
IncHigh	Higher level income	-10.04	9.11	0.27	-27.9	7.83
copen	Lives in Copenhagen area	-11.02	9.97	0.27	-30.56	8.52
city	Lives in larger city of 20,000 or more	-1.8	11.81	0.88	-24.94	21.35
σ_u	Standard deviation of individual effect	15.58	1.43	0	12.79	18.38
σ_e	Standard deviation of residual	10.43	0.62	0	9.22	11.64

Notes: Log-likelihood value is -708.7; Wald test for null hypothesis that all coefficients are zero has a χ^2 value of 58.3 with 21 degrees of freedom, implying a *p*-value of less than 0.0001; fraction of the total error variance due to random individual effects is estimated to be 0.69, with a standard error of 0.048.

Legend: Most variables have self-evident definitions, or are defined under Table 3.

Table 5: Statistical Model of Risk Aversion Responses in the United States

Random-effects interval regression,
with the final CRRA interval chosen by the subject as the dependent variable.

N= 441, based on 115 subjects.

Variable	Description	Estimate	Standard Error	*p*-value	Lower 95% Confidence Interval	Upper 95% Confidence Interval
Constant		0.43	0.32	0.17	-0.19	1.06
impl	iMPL format	0.08	0.14	0.56	-0.19	0.35
skewLO	SkewLO frame	-0.07	0.06	0.23	-0.19	0.05
skewHI	SkewHI frame	0.09	0.06	0.14	-0.03	0.20
endowment	Random initial endowment	-0.02	0.02	0.38	-0.06	0.02
second	Risk aversion second task	0.19	0.13	0.14	-0.06	0.44
Task2	Second task	0.03	0.07	0.69	-0.12	0.17
Task3	Third task	0.01	0.07	0.93	-0.13	0.14
Task4	Fourth task	0.05	0.06	0.37	-0.06	0.17
female	Female	-0.07	0.11	0.54	-0.28	0.15
single	Lives alone	-0.02	0.13	0.89	-0.27	0.24
nhhd	Number in household	0.08	0.05	0.14	-0.03	0.18
owner	Owns home or apartment	0.06	0.12	0.61	-0.18	0.31
student	Student	-0.08	0.15	0.60	-0.37	0.22
skilled	Some post-secondary education	0.09	0.12	0.46	-0.14	0.32
IncLow	Lower level income	0.05	0.15	0.74	-0.25	0.35
IncHigh	Higher level income	0.08	0.15	0.58	-0.21	0.38
city	Lives in larger city of 20,000 or more	0.24	0.16	0.14	-0.07	0.55
σ_u	Standard deviation of individual effect	0.51	0.04	0.00	0.42	0.59
σ_e	Standard deviation of residual	0.41	0.02	0.00	0.38	0.44

Notes: Log-likelihood value is -1149106; Wald test for null hypothesis that all coefficients are zero has a χ^2 value of 18.72 with 17 degrees of freedom, implying a *p*-value of 0.3447; fraction of the total error variance due to random individual effects is estimated to be .60, with a standard error of .0472.

Legend: Most variables have self-evident definitions, or are defined under Table 3. The variable "second" indicates subjects facing the risk task after participating in a willingness to pay task.

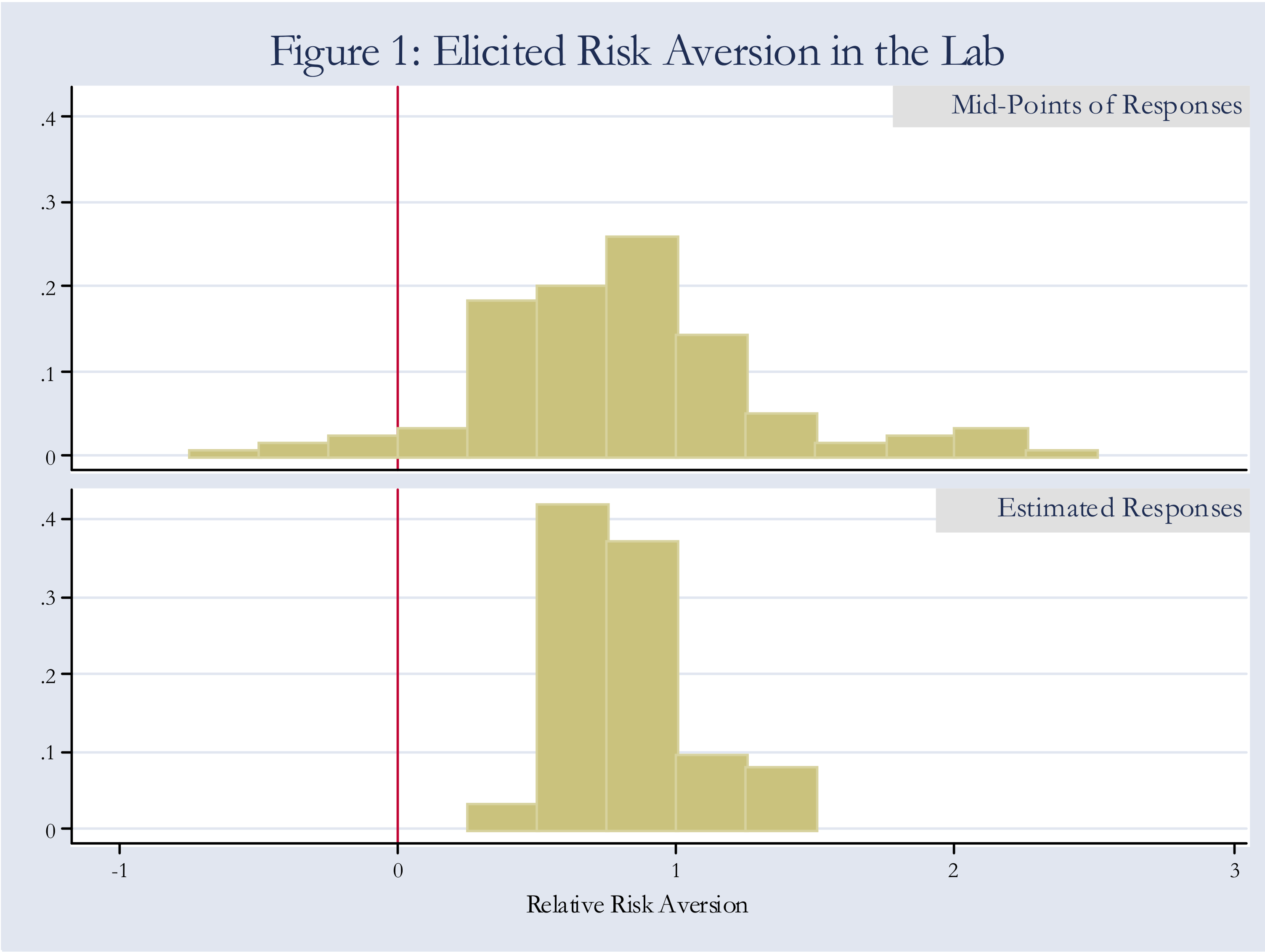
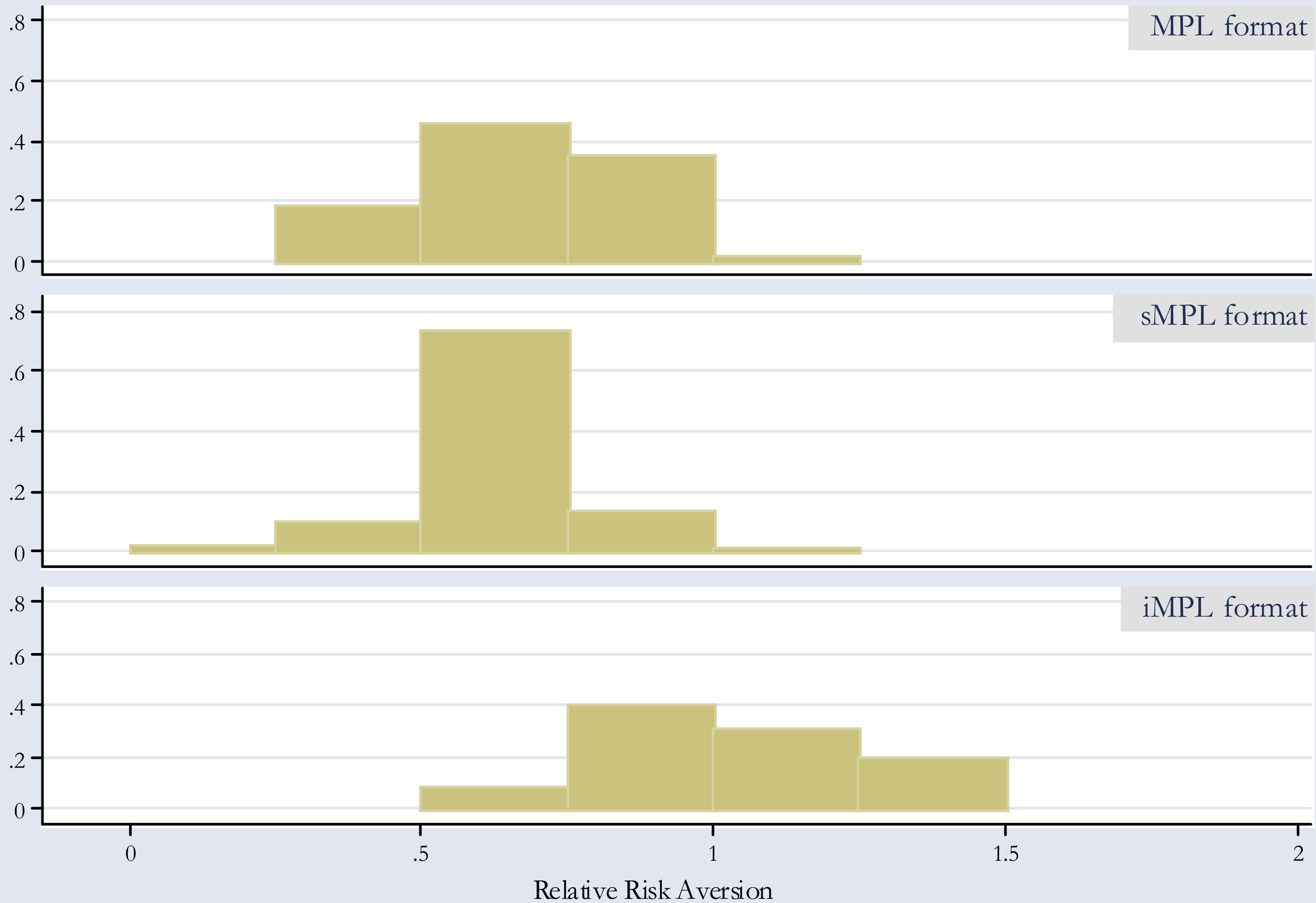


Figure 1: Elicited Risk Aversion in the Lab
Mid-Points of Responses
Estimated Responses
0
.1
.2
.3
.4
-1
0
1
2
3
Relative Risk Aversion

Figure 2: Estimated Risk Aversion and Format



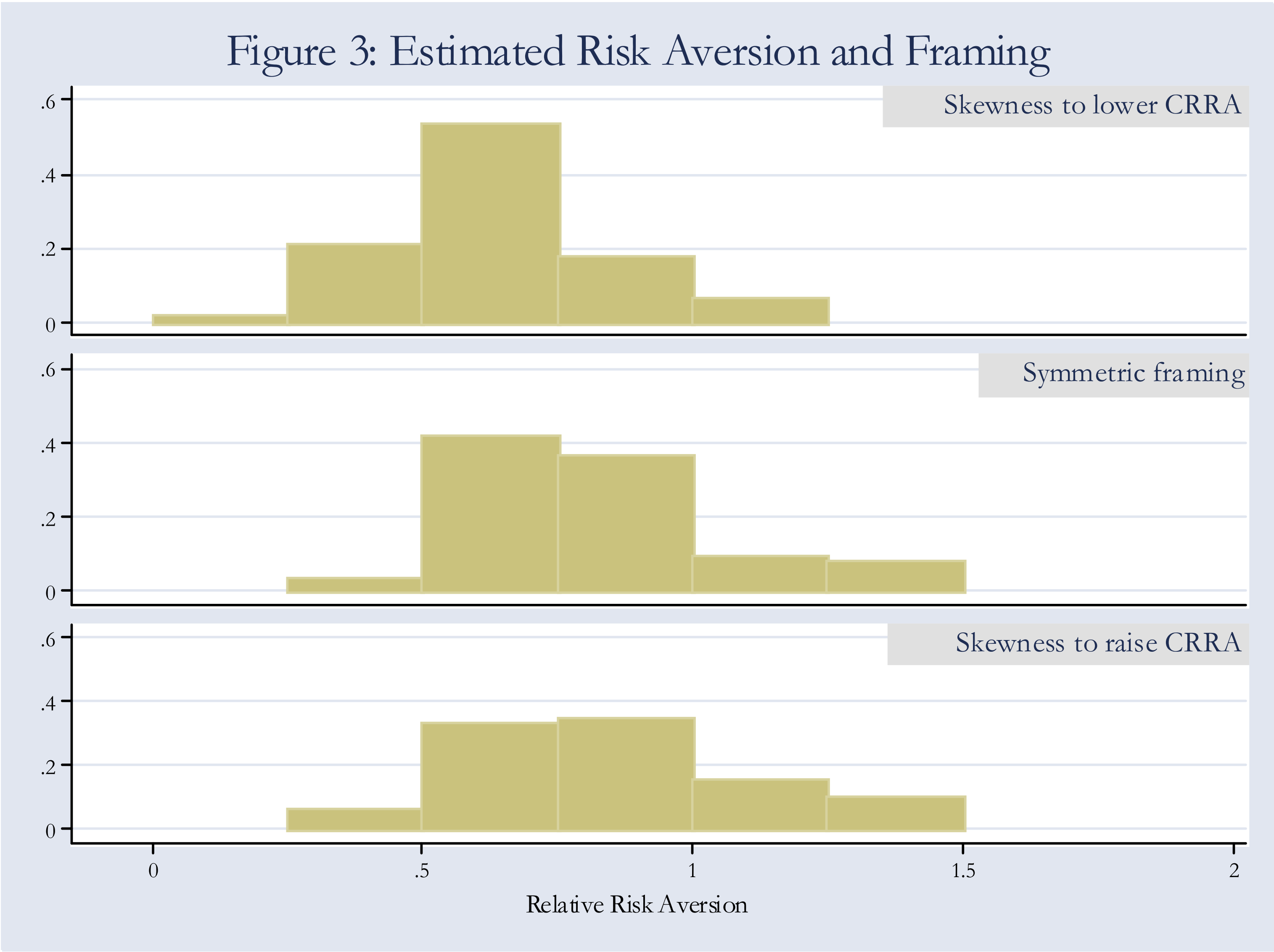


Figure 3: Estimated Risk Aversion and Framing
Skewness to lower CRRA
Symmetric framing
Skewness to raise CRRA
.6
.4
.2
0
0
.5
1
1.5
2
Relative Risk Aversion

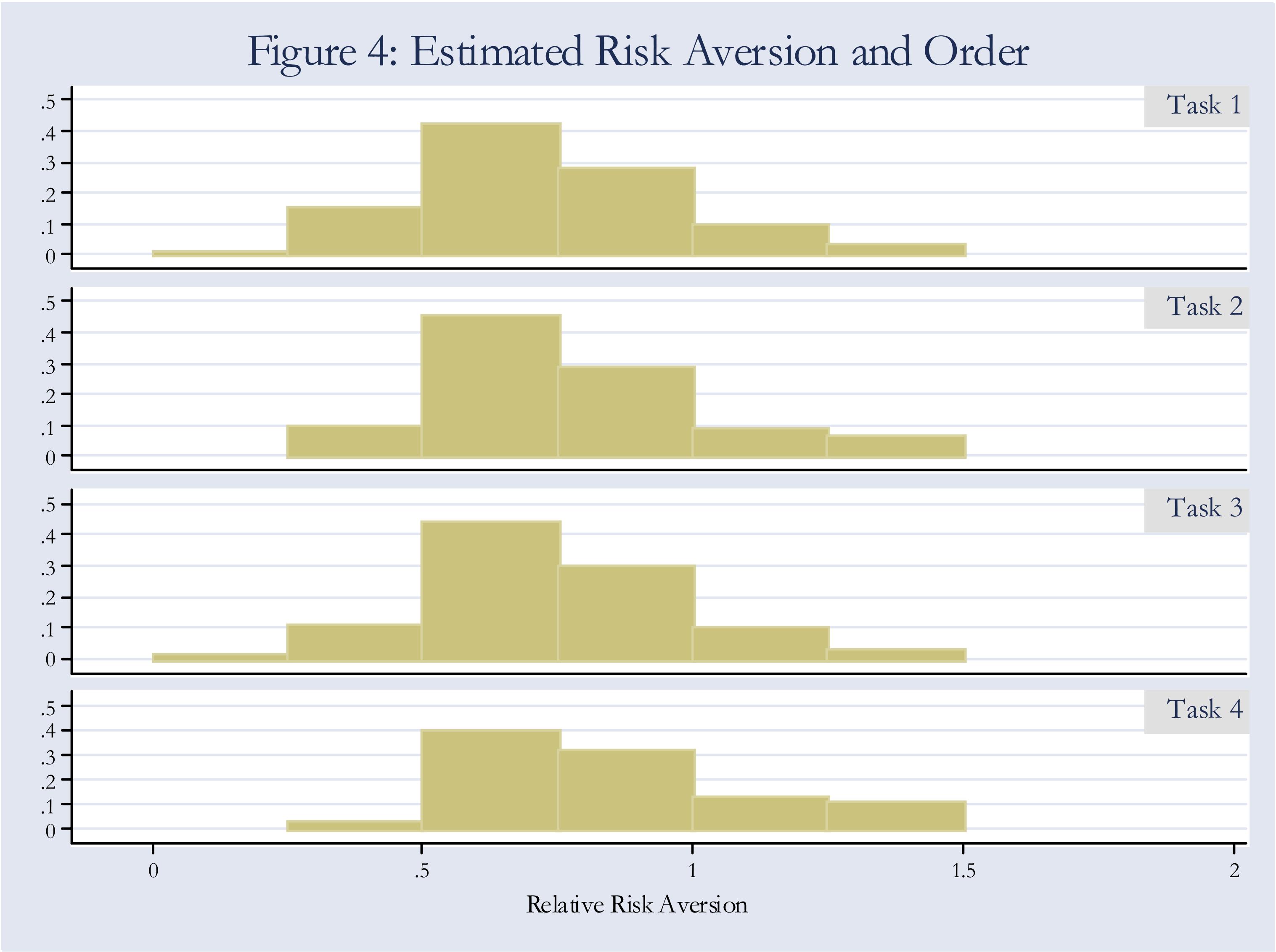


Figure 4: Estimated Risk Aversion and Order
Task 1
Task 2
Task 3
Task 4
.5
.4
.3
.2
.1
0
0
.5
1
1.5
2
Relative Risk Aversion

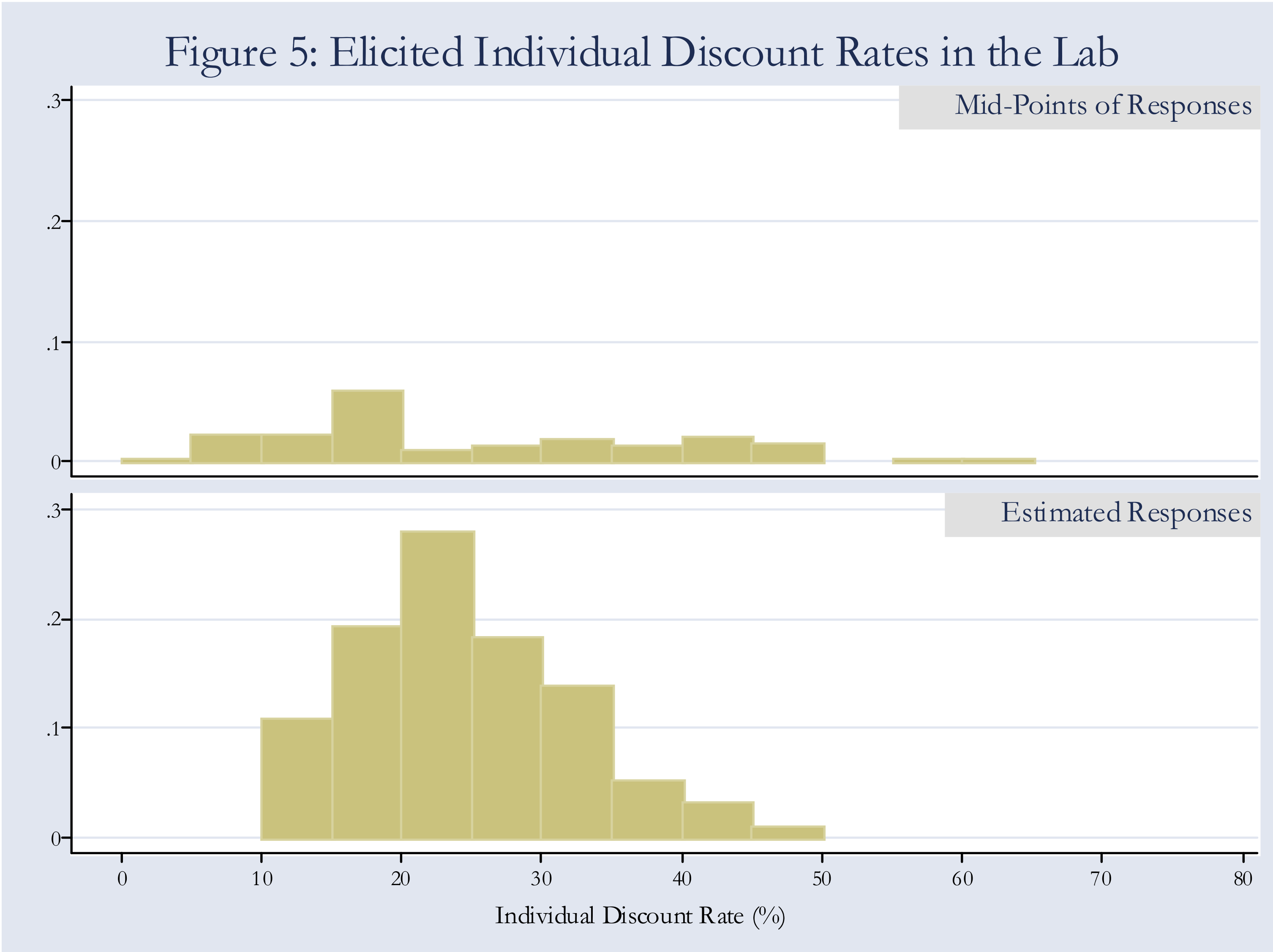


Figure 5: Elicited Individual Discount Rates in the Lab
Mid-Points of Responses
Estimated Responses
.3
.2
.1
0
0
10
20
30
40
50
60
70
80
Individual Discount Rate (%)

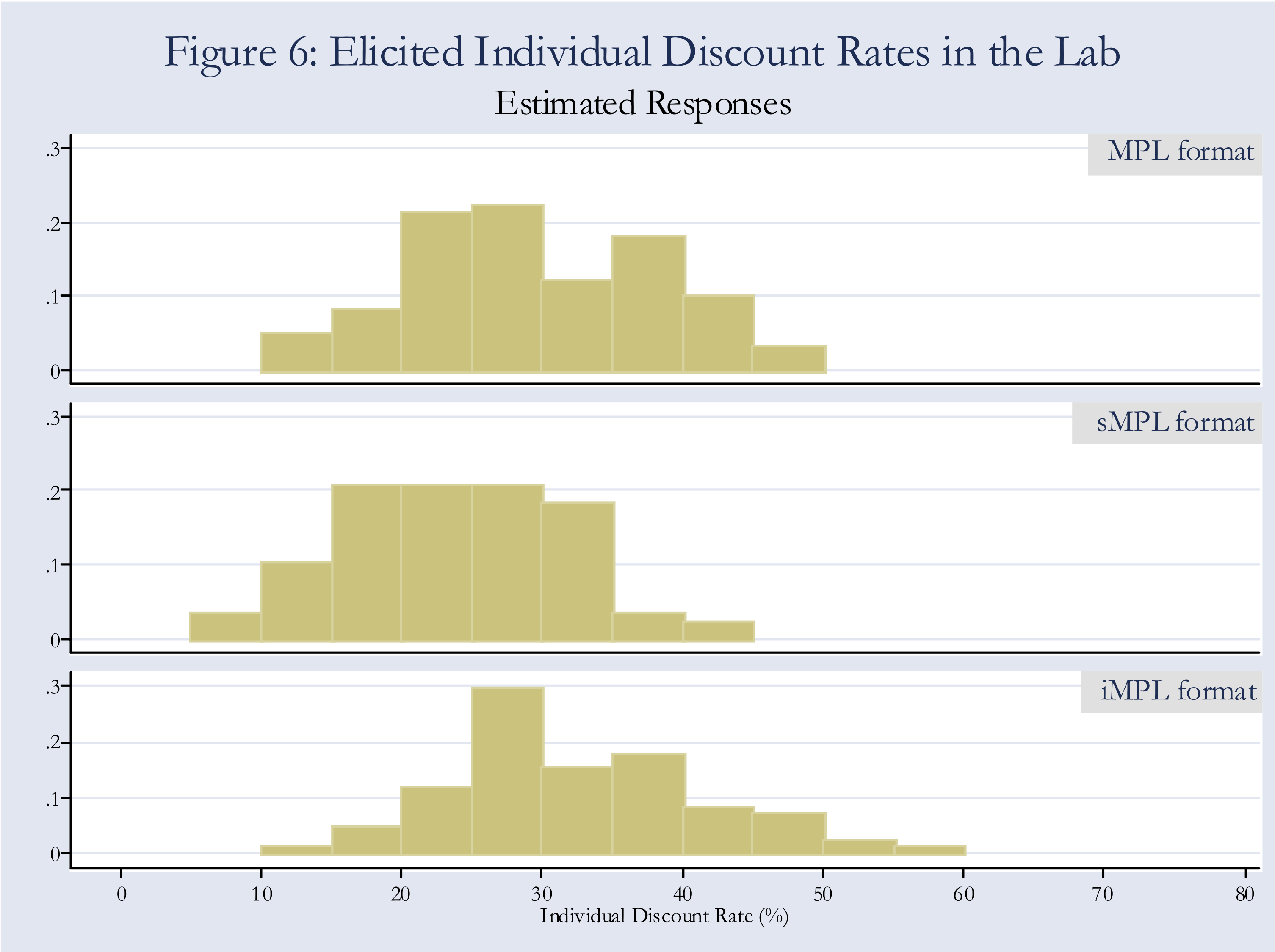


Figure 6: Elicited Individual Discount Rates in the Lab
Estimated Responses
MPL format
sMPL format
iMPL format
0
.1
.2
.3
0
10
20
30
40
50
60
70
80
Individual Discount Rate (%)

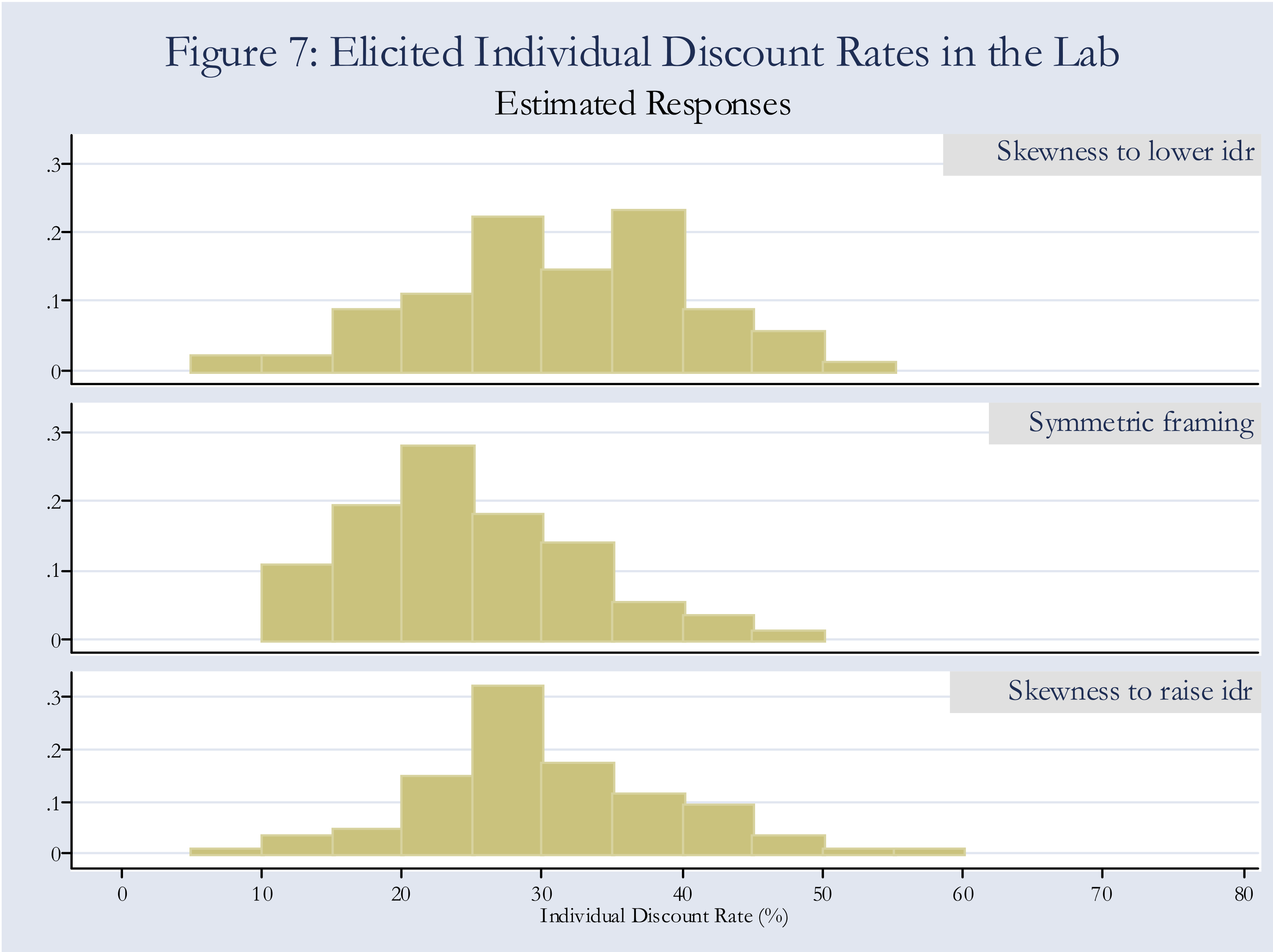


Figure 7: Elicited Individual Discount Rates in the Lab
Estimated Responses
Skewness to lower idr
Symmetric framing
Skewness to raise idr
0
.1
.2
.3
0
10
20
30
40
50
60
70
80
Individual Discount Rate (%)

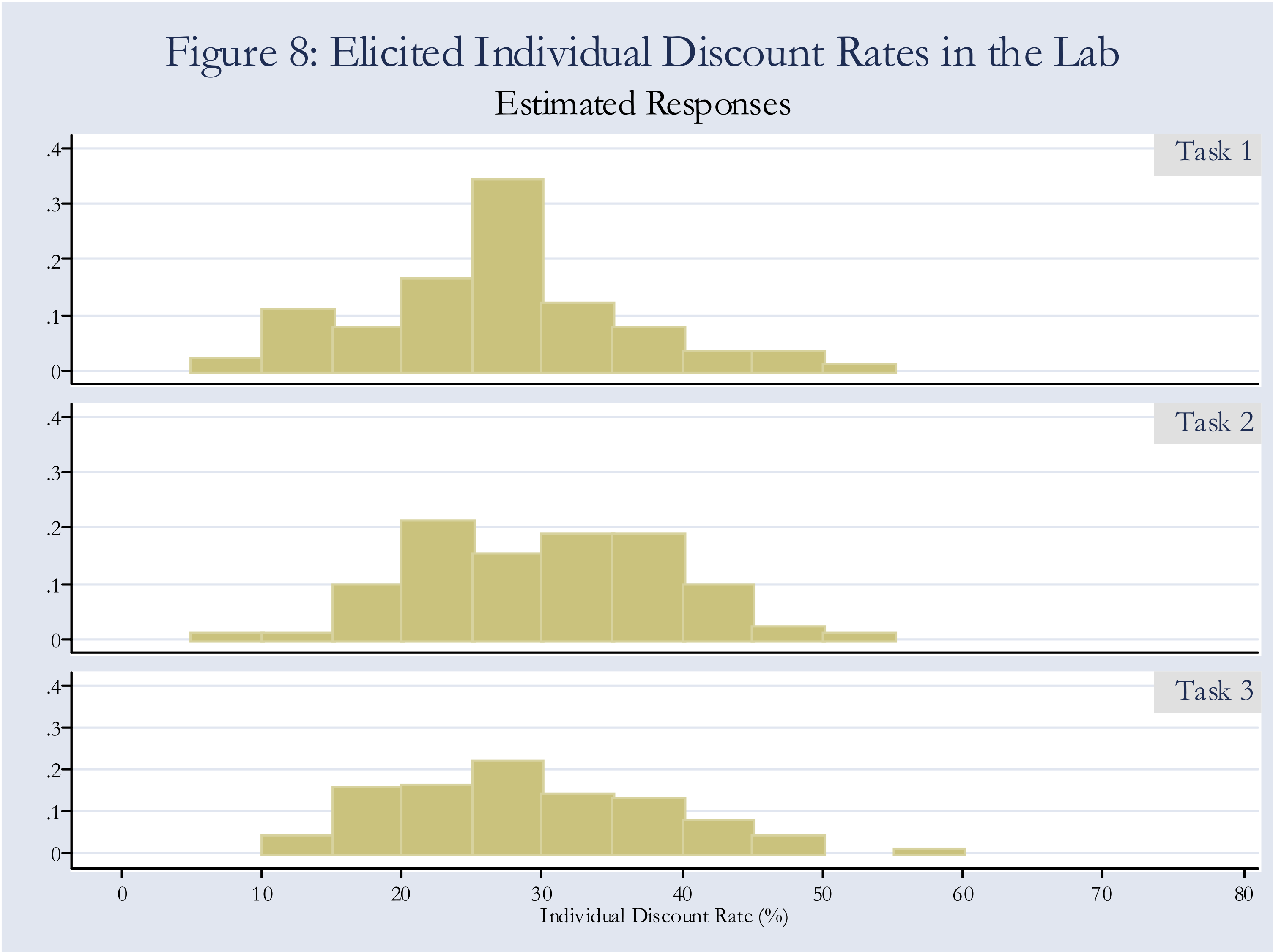


Figure 8: Elicited Individual Discount Rates in the Lab
Estimated Responses
Task 1
Task 2
Task 3
.4
.3
.2
.1
0
0
10
20
30
40
50
60
70
80
Individual Discount Rate (%)

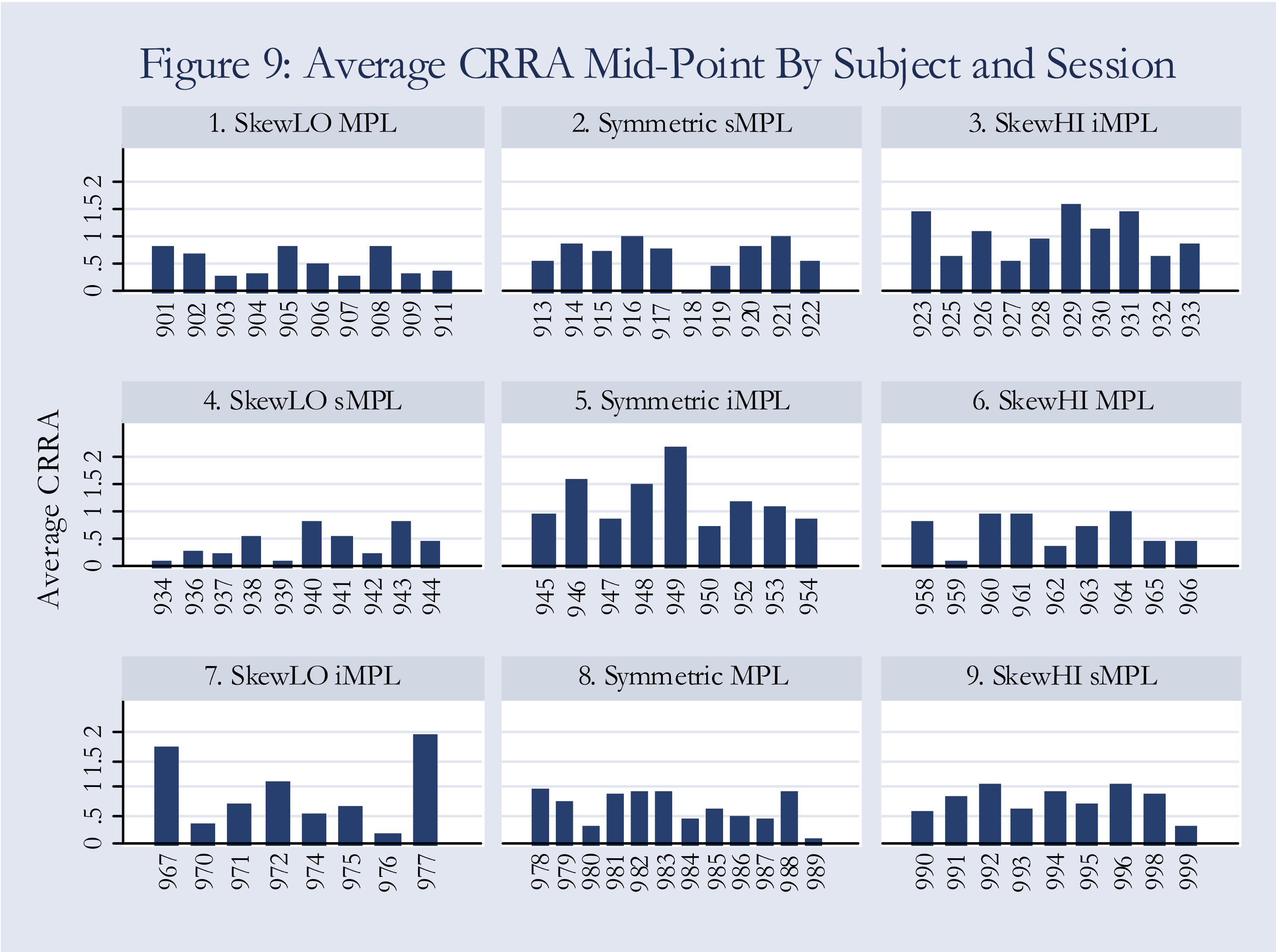


Figure 9: Average CRRA Mid-Point By Subject and Session
1. SkewLO MPL
2. Symmetric sMPL
3. SkewHI iMPL
4. SkewLO sMPL
5. Symmetric iMPL
6. SkewHI MPL
7. SkewLO iMPL
8. Symmetric MPL
9. SkewHI sMPL
Average CRRA
0 .5 1 1.5 2
901 902 903 904 905 906 907 908 909 911
913 914 915 916 917 918 919 920 921 922
923 925 926 927 928 929 930 931 932 933
934 936 937 938 939 940 941 942 943 944
945 946 947 948 949 950 952 953 954
958 959 960 961 962 963 964 965 966
967 970 971 972 974 975 976 977
978 979 980 981 982 983 984 985 986 987 988 989
990 991 992 993 994 995 996 998 999

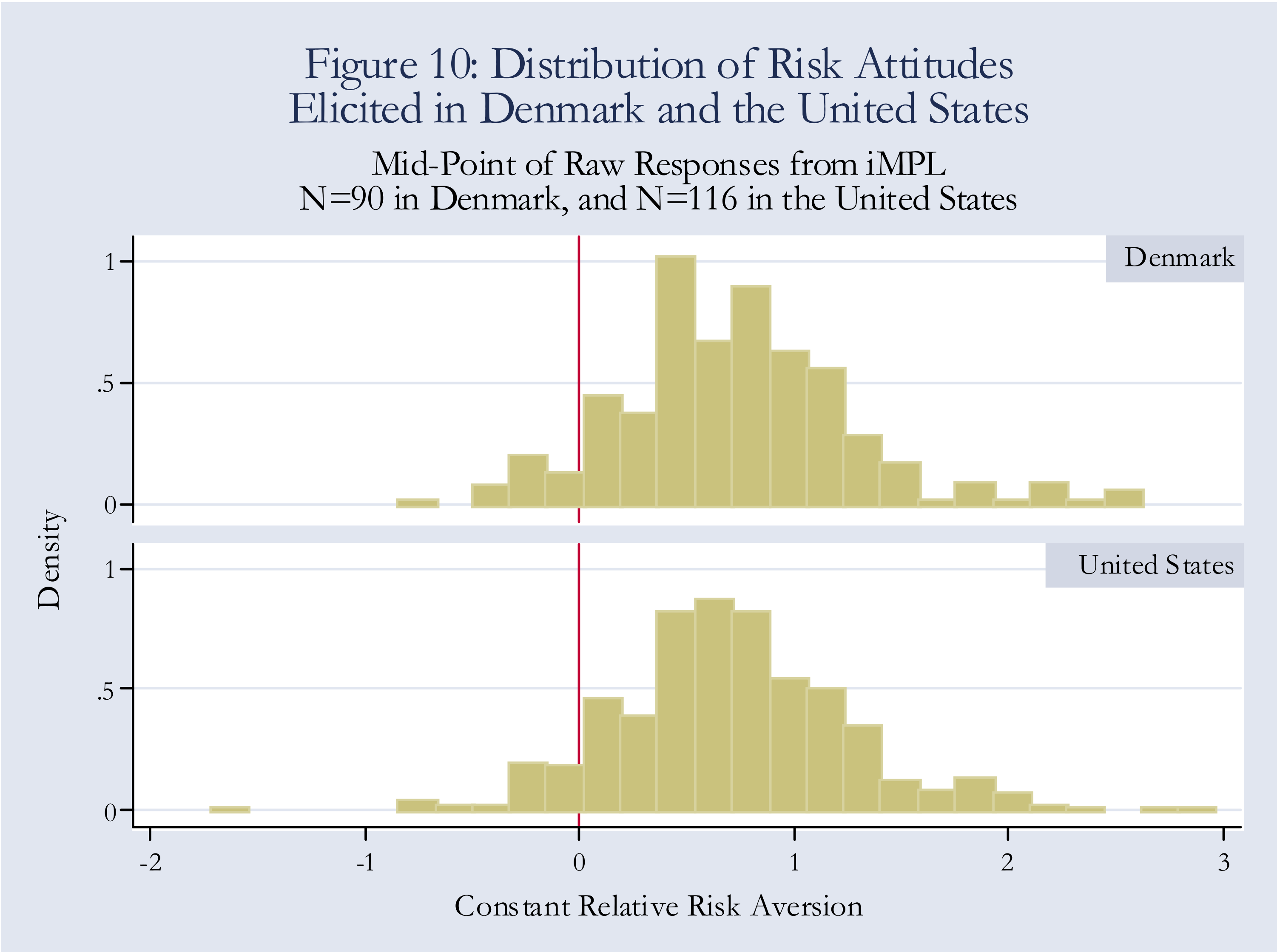


Figure 10: Distribution of Risk Attitudes
Elicited in Denmark and the United States
Mid-Point of Raw Responses from iMPL
N=90 in Denmark, and N=116 in the United States
Denmark
United States
Density
1
.5
0
1
.5
0
-2
-1
0
1
2
3
Constant Relative Risk Aversion

References

Beck, John H., "An Experimental Test of Preferences for the Distribution of Income and Individual Risk Aversion," *Eastern Economic Journal*, 20(2), Spring 1994, 131-145.

Binswanger, Hans P., "Attitudes Toward Risk: Experimental Measurement in Rural India," *American Journal of Agricultural Economics*, 62, August 1980, 395-407.

Binswanger, Hans P., "Attitudes Toward Risk: Theoretical Implications of an Experiment in Rural India," *Economic Journal*, 91, December 1981, 867-890.

Botelho, Anabela; Harrison, Glenn W.; Hirsch, Marc A., and Rutström, Elisabet E., "Bargaining Behavior, Demographics and Nationality: What Can the Experimental Evidence Show?" 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).

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

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

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

Davis, Douglas D., and Holt, Charles A., *Experimental Economics* (Princeton, NJ: Princeton University Press, 1993).

Eckel, Catherine C., and Grossman, Philip J., "Forecasting Risk Attitudes: An Experimental Study of Actual and Forecast Risk Attitudes of Women and Men," *Unpublished Manuscript*, Department of Economics, Virginia Tech, October 2002.

Friedman, Daniel and Sunder, Shyam, *Experimental Methods: A Primer for Economists* (New York: Cambridge University Press, 1994).

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

Gonzalez, Richard, and Wu, George, "On the Shape of the Probability Weighting Function," *Cognitive Psychology*, 38, 1999, 129-166.

Harrison, Glenn W., "Theory and Misbehavior of First-Price Auctions: Reply," *American Economic Review*, 82, December 1992, 1426-1443.

Harrison, Glenn W., "Experimental Evidence on Alternative Environmental Valuation Methods," *Environmental & Resource Economics*, 2004 forthcoming.

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

Harrison, Glenn W.; Johnson, Eric; McInnes, Melayne M., and Rutström, E. Elisabet, "Risk Aversion and Incentive Effects: Comment," *Working Paper 03-19*, Department of Economics, College of

Business Administration, University of Central Florida, 2003a; forthcoming, *American Economic Review*.

Harrison, Glenn W.; Johnson, Eric; McInnes, Melayne M., and Rutström, E. Elisabet, "Risk Aversion and Individual Choice in the Laboratory: A Reconsideration," *Working Paper 03-18*, Department of Economics, College of Business Administration, University of Central Florida, 2003b.

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

Hey, John D., *Experiments in Economics* (Cambridge, UK: Blackwell, 1991).

Hoffman, Elizabeth; Menkhaus, Dale J.; Chakravarti, Dipinkar; Field, Ray A., and Whipple, Glen D., "Using Laboratory Experimental Auctions in Marketing Research: A Case Study of New Packaging for Fresh Beef," *Marketing Science*, 12(3), Summer 1993, 318-338.

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

Kagel, John, H., and Roth, Alvin E. (eds.), *The Handbook of Experimental Economics* (Princeton: Princeton University Press, 1995).

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

Kirby, Kris N. and Maraković, Nino N., "Delay-discounting Probabilistic Rewards: Rates Decrease as Amounts Increase," *Psychonomic Bulletin & Review*, 1996, 3:1, 100-104.

Kirby, Kris N.; Petry, Nancy M., and Bickel, Warren K., "Heroin Addicts Have Higher Discount Rates for Delayed Rewards than Non-drug-using Controls," *Journal of Experimental Psychology: General*, 1999, 128:1, 78-87.

Laury, Susan K., and Holt, Charles A., "Further Reflections on Prospect Theory," *Working Paper*, Department of Economics, Georgia State University, 2002.

Laury, Susan K., and Holt, Charles A., "Payoff Scale Effects and Risk Preference Under Real and Hypothetical Conditions," in C.R. Plott and V.L. Smith (eds.), *Handbook of Experimental Economics Results* (North-Holland: Amsterdam, forthcoming 2004).

Miller, Louis; Meyer, David Edward, and Lanzetta, John T., "Choice Among Equal Expected Value Alternatives: Sequential Effects of Winning Probability Level on Risk Preferences," *Journal of Experimental Psychology*, 79(3), 1969, 419-423.

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

Murnighan, J. Keith; Roth, Alvin E., and Shoumaker, Francoise, "Risk Aversion and Bargaining: Some Preliminary Results," *European Economic Review*, 31, 1987, 265-271.

Murnighan, J. Keith; Roth, Alvin E., and Shoumaker, Francoise, "Risk Aversion in Bargaining: An Experimental Study," *Journal of Risk and Uncertainty*, 1(1), March 1988, 101-124.

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

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

Shogren, Jason F. "Experimental Methods and Valuation," in K-G. Mäler and J. Vincent (eds.), *Handbook of Environmental Economics. Volume2: Valuing Environmental Changes* (Amsterdam: North-Holland, 2004).