Eliciting Risk Preferences: When is Simple Better?

Chetan Dave	Catherine Eckel	Cathleen Johnson	Christian Rojas
University of Texas at Dallas	University of Texas at Dallas	University of Arizona	University of Massachusetts Amherst
cdave@utdallas.edu	eckelc@utdallas.edu	cjohnson@eller.arizona.edu	rojas@resecon.umass.edu

August, 2007

Abstract

We study the estimation of risk aversion preferences with experimental data. We focus on the trade-offs that arise when choosing between two different elicitation methods that have different degrees of difficulty for subjects. We analyze how and when the simpler, but coarser, elicitation method may be preferred to the more complex, but finer, one. Subjects in the more complex task exhibit noisier behavior, especially if their mathematical ability is low. While the advantage of the more complex measure is an *overall* superior predictive accuracy, the simpler task generates less noise and similar predictive accuracy for subjects with *low mathematical ability*. Furthermore, in a retest of the experiment, we find that subjects appear to be slightly less time-consistent in the more complex task. We also find important heterogeneity in preferences and provide methodological suggestions for future work.

Things should be made as simple as possible, but not simpler. ----Albert Einstein

1. Introduction

In experimental decision making, individuals are assumed to reveal their preferences as long as the performed tasks have salient payoff consequences. A number of methods can be employed to elicit such preferences and the resulting data can be used to estimate the parameters of a utility function. Since many economic models of individual and family decision making rely on a parameterized utility function to make precise predictions, being able to effectively measure individual-specific risk preferences is important for predicting and understanding outcomes and preference heterogeneity across individuals.

In this paper we explore several issues involved in estimating the underlying risk preferences of individuals using experimental data. First, we study the trade-offs that arise when choosing between two different elicitation methods that have different degrees of difficulty for subjects. We place a particular emphasis on the population where task complexity may be more critical: people with limited mathematical ability. A brief description of the data should make this objective clearer. As part of a larger study, we collected observations on two experimental measures[1] of risk preferences for nearly 900 adults, together with an extensive survey and a widely-validated measure of mathematical ability. The experimental measures differ in that one (developed by Holt and Laury, 2002 – HL hereafter) is more complex, involving ten decisions between gambles with probabilities ranging from .1 to .9, and allows categorization of decision makers into 10 risk categories, while the other (developed by Eckel and Grossman, 2002, 2007 - EG hereafter) is simpler, involving a single choice among 6 gambles, all with .5 probability of winning a higher prize, but only allows categorization of decision makers into 5 risk categories.

[1] We use the terms "elicitation methods" and "experimental measures" interchangeably throughout the text.

Our empirical approach follows Holt and Laury (2002). We estimate a constant relative risk aversion (CRRA) utility function that allows for a parameter that captures noisy behavior. A significant part of our analysis compares the estimated noise parameter between the two risk elicitation methods, so it is important to describe the interpretation of this parameter in this particular specification. A small noise parameter indicates that subjects' decisions conform closely to the predictions of expected utility theory, whereas a large noise parameter is considered as evidence of a high degree of randomness in subjects' decisions. The advantage of the cognitively more difficult measure (HL) is that it produces a finer categorization; as a result, estimates from the HL measure have higher predictive accuracy (as measured by the percentage of choices that are correctly predicted by the estimates). However, complexity comes at a cost: subjects' decisions exhibit larger noise, especially among the low math ability population. Conversely, the disadvantage of the simpler measure is that it produces a coarser categorization (and thus lower predictive accuracy), but decisions are substantially less noisy. A central result of this paper is that the simpler measure appears to be unambiguously superior for low math ability subjects as it generates smaller noise and equal predictive accuracy than the complex measure.

Besides the trade-offs between noise and predictive accuracy, we also investigate two other important issues: time stability of estimates and heterogeneity in preferences. A subset of the subjects participated in a retest of the protocol six months after the initial experiment; this feature allows us to study the time-stability of estimates from the two methods by comparing the estimated parameters of this subpopulation between these two time periods. Results suggest that the simpler experimental measure may have slightly better retest stability.

Heterogeneity in preferences is analyzed by incorporating observable individual characteristics in the risk preference parameter.[2] In addition to the math ability measure discussed

[2] The noise parameter is also specified as a function of individual characteristics.

earlier, these characteristics include gender, income and age. Considerable evidence points to greater risk aversion among women (see Eckel 2007, Eckel and Grossman, 2007, and the references therein), and it is plausible that risk aversion may diminish at higher income levels. Conversely, younger persons (especially males) may tend to be more willing to take on risks (e.g., Dohmen, et al., 2006). Results confirm earlier studies' results of higher risk aversion for females; this effect is strongly is present in our data as it appears in all specifications and in both risk elicitation methods. Also, EG (but not HL) shows greater measured risk aversion for subjects with low math ability. Interestingly, the more complex measure shows a greater level of overall risk aversion.[3] It is not immediately obvious how to interpret the latter two results as it is unclear whether the difference in estimates across methods is an artifact of noisier behavior (or "mistakes") in decision making. We discuss alternative interpretations of these findings in the results and conclusion sections.

Our overall assessment of our results is that the type of measurement technique may be critical in recovering reliable estimates of risk preferences. Economists often neglect the importance of how a task is presented to subjects – in particular, the difficulty of the task– but focus instead the theoretical properties of such tasks. Complexity seems to be an important issue when designing an appropriate method, especially for low math ability populations. Thus, care should be taken to develop experimental measures of preferences that have desirable properties for the population under study.

In Section 2 we provide a brief description of the experimental data and in Section 3 we describe the methodology employed. Section 4 reports model estimates and Section 5 concludes with the main results.

2. Description of Experiments and Data

2.1 Implementation

[3] This difference is smaller when we control for heterogeneity in the noise parameter.

From May 2002 to March 2003, 881 Canadian residents, ranging in age from 18 to 54 years, participated in 102 experimental sessions. This sample was drawn from both urban and non-urban sites across Canada and was made up mainly of people who were engaged in the labor force. The sample includes three age groups: youth between the ages of 18 and 24 years, including 80 high school students and slightly older participants who were not full-time students; subjects aged 25-44 who had some labor-force attachment, including both employed or unemployed samples; and a small number of additional subjects aged 45-55. Table 1 gives sample details.

Unemployed participants were recruited by local Human Resource Centre of Canada (HRCC) staff and recruitment pamphlets placed in HRCC waiting areas. Other participants were targeted through newspaper advertising in popular daily newspapers, TV and radio announcements on community stations. Site visits to local businesses and community organizations also proved to be effective recruiting tools, especially for younger participants. Recruitment materials included information about the project, time commitment, a show-up fee (CAD$20), the potential for more earnings, and confidentiality of responses and participation. Participants volunteered by calling a toll-free number or signing up through a web page. Prior to scheduling, they completed a short demographic questionnaire to determine eligibility for participation according to the sample design.

The experimental sessions were conducted in controlled environments including classrooms, boardrooms and hotel conference facilities. Upon arrival, the experimenter reminded participants of the confidentiality of the data, and provided participants with appropriate details of the potential earnings, including the possibility of cash payments. The maximum number of participants in any session was 30.

Care was taken to ensure that subjects understood the decisions they were to make. Because these decisions were unfamiliar, thirteen practice examples, one of each kind of decision, were demonstrated to ensure that they understood the nature of the decisions and how payment was linked to their choices.

Subjects completed a series of experimental decisions involving choices between two or more alternatives. After all decisions were made, one decision was selected at random (from all the experimental decisions) for payment using a bingo ball cage where each decision number was matched with one corresponding bingo ball number.[4] Each had an equal probability of being selected, making decisions independent of each other. Subjects were paid in private, and average earnings were CAD$165. Each session took between 1.5 and 3 hours to complete from instruction to payoff.

2.2 Risk Preferences

To measure risk preferences, participants in the experiment completed two different sets of decisions involving choices among cash gambles, as shown in the Appendix. Subjects' choices and details of the gamble sets are summarized in Tables 2 and 3. In the first experimental measure, subjects complete a task developed by Eckel and Grossman (2002, 2007) where risk attitudes are measured by having subjects choose from among six possible gambles the one they would most prefer to play.[5] To determine payment, the subject plays the gamble chosen by rolling a die.

Table 2 and Figure 1 illustrate the gambles in the choice set. All the gambles involve a 50/50 chance of a low or high payoff. The range of gambles includes a safe alternative involving a sure payoff with zero variance. The gambles increase in both expected return and risk (standard deviation) moving from Gamble 1 to 5. Gamble 6 involves only an increase in variance with the

[4] There were 13 types of decisions, and 100 decisions in total.

[5] This approach to eliciting risk preferences is similar to one developed by Binswanger (1980, 1981), which also uses a set of 50/50 gambles that vary in risk and expected payoff. The primary differences between this protocol and Binswanger's are his inclusion of dominated gambles, nonlinear relationship between risk and return, and presentation of tasks in a pairwise format. Ours was developed independently. Recent implementations of Binswanger's protocol adopt a format very similar to our own, and have been used successfully in the field with less literate populations (e.g. Barr, 2003).

same expected return as gamble 5. More risk-averse subjects would choose lower-risk, lower-return gambles; risk-neutral subjects would choose gamble 5 or 6, which have the highest rate of return; risk-seeking subjects would choose gamble 6. This task was designed to be as simple as possible, while retaining a reasonable range of risky choices, and takes only a few minutes to explain and implement.

Table 2 also includes of coefficients of relative risk aversion implied by each possible choice under the assumption of constant relative risk aversion (CRRA), shown as 'r' in the table. In each case, the range was calculated by comparing each gamble to its neighbors and calculating the value of r that generates the same utility level for the payoffs associated with each adjacent gamble. A person choosing Gamble 3, for example, would have a coefficient of relative risk aversion in the range 0.71-1.16: a person with $r = 0.71$ would be just indifferent between gambles 3 and 4, and a person with $r = 1.16$ is just indifferent between gambles 2 and 3. A choice of gamble 6 implies risk-seeking, with $r < 0$. The distribution of choices shown here is very similar to samples of university students with stakes 1/3 to 1/2 this size (e.g., Eckel and Grossman 2007).

Table 3 provides the r ranges and the frequency of safe choices for the risk aversion measure developed by Holt and Laury (2002). This widely-used experimental measure involves a set of ten binary choices between a high risk gamble and a low risk gamble. The two gambles have the same probabilities but different low and high payoffs, making them relatively easy to compare. Most subjects quickly see that they should prefer Gamble A at the top of the decision sheet, and Gamble B at the bottom, implying some switch point in between. The switch point determines the number of safe choices and, in turn, the risk aversion parameter range. A risk neutral subject would switch between decisions 4 and 5, making 4 safe decisions. Subjects in this sample tend to be risk averse, making more than 4 safe decisions. Using the number of safe choices as an aggregate measure is not a fully accurate summary of the distribution of choices because some subjects (8.5% of the sample) make *inconsistent* decisions, either by switching more than once or by making "backwards" choices (switching in the other direction). The last column of the table removes these subjects; the resulting distribution is slightly more risk averse, overall. This distribution of choices is very similar to those reported in Holt and Laury (2007) for university students making decisions at this stakes level (equivalent to their 20X Real treatment).

Comparing the two methods, the EG measure involves fewer (and simpler) gambles and a single choice, while the HL measure involves more complex gambles and more choices. Part of our purpose in including both tasks was to compare the two elicitation methods, and examine the tradeoff involved in having a coarser screen but an easier set of decisions for subjects to make. Our experience in the field and in the lab is that this difference in difficulty may not to be an issue for university students, but less well-educated subjects may find the EG method easier to understand.

We conjecture that important tradeoffs may be involved in developing experimental tasks to measure risk attitudes. Expected utility theory suggests a straightforward approach to measuring risk: it implies that any task that explores the curvature of the function will produce data that can be used to parameterize a utility function, generating, for example, a CRRA coefficient of relative risk aversion. If one is willing to assume the existence of a well-behaved utility function, and precise decision making skills on the part of subjects, this approach is sensible. However, some researchers have found that subjects' preferences are inconsistent across similar tasks (e.g., Slovic 1964; Isaac and James, 2000; Berg, et al., 2005; Peters, et al., 2006). This suggests that the possibility that subjects make decisions with some degree of error should be taken into account.

2.3 Ability

Arguably, a subject's mathematical ability may interfere with his ability to understand and choose among risky alternatives, and hence with the experimentalist's objective to infer the degree risk aversion. The study included a questionnaire that assesses mathematical ability (numeracy); we use it here to explore the relationship between noisy behavior and ability. This numeracy assessment is a subcomponent of the Educational Testing Service's Adult Literacy and Lifeskills Survey (ALLS) and provides a proxy for participants' readiness to learn and to engage in educational activities. This is a widely used and validated literacy test; the score on the test is a strong predictor of income earning ability (Statistics Canada and OECD, 2003). The numeracy assessment consists of 31 problems involving the use of mathematics in real-life situations. The results of this assessment provide a gauge of an individual's competencies. Subjects completed this unpaid task following completion of the experimental decisions. The distribution of scores in this study is very similar to those for Canadians with post-secondary education (Statistics Canada and OECD, 2003).[6]

As an initial indication of the potential importance of math ability for accurate measurement of risk attitudes, we calculated the number of people making inconsistent choices in HL by ability level. Figure 2 shows that the fraction of inconsistent subjects is much larger in the low math-ability population (subjects who score less than one standard deviation below the sample mean of math ability scores) than in the remaining population.

2.4 Coding Choice Data

To proceed with the econometric analysis, we convert choices in both elicitation methods to a binary format. In each of the 10 decisions of the HL method, the risky choice (B) takes a value of 1, and the less risky choice (A) takes a value of 0. Subjects in the EG method make only one choice; hence EG choices are not directly comparable to HL choices. To make data in both methods as comparable as possible, we transform the EG choice into data that has a format similar to that of

[6] 5, 25, 50, 75 and 95 percentile scores for our sample (and for Canada population) are: 34 (36), 44 (44), 50 (49), 57 (55) and 65 (62), respectively.

HL. First we convert the six gambles into the five (HL equivalent) hypothetical binary decisions that result from comparing adjacent gambles:[7]

Decision 1: Gamble 5 v. Gamble 6
Decision 2: Gamble 4 v. Gamble 5
Decision 3: Gamble 3 v. Gamble 4
Decision 4: Gamble 2 v. Gamble 3
Decision 5: Gamble 1 v. Gamble 2

We have arranged these decisions to parallel the order of the decisions in HL: the more risky gamble is located on the right, and decision 5 corresponds to subjects with the highest level of risk aversion. To illustrate the coding procedure, consider a subject chooses Gamble 4 in the EG method. This implies a coding of 0, 0, 1, 1 and 1, for each of the five hypothetical decisions, respectively. Note that, as opposed to the HL method, this coding of the EG data does not allow subjects to be inconsistent (switching back and forth, or switching "backwards"). We address this issue in the estimation by analyzing the sensitivity of our results when inconsistent subjects are excluded.

To illustrate that coding the EG data as pair-wise comparisons between adjacent gambles is not an unreasonable depiction, consider Figure 1. Risk-averse indifference curves in this space would be upward sloping, with a tangency on the upward sloping line representing the subject's preferred choice. Suppose a subject chooses gamble 3, with a tangency at that point. For well-behaved utility functions, this implies that gamble 2 is preferred to 1, 3 preferred to 2 and 3 is preferred to 4, 4 to 5, and 5 to 6.

2.5 Other Variables and Additional Data Details

We construct several variables to explore the issue of whether estimated parameters vary by experimental measure and by subjects' characteristics. A dichotomous variable, denoted HL, takes a value of 1 if the choice corresponds to the HL measure and zero otherwise. Variables for subjects'

[7] Engle-Warnick et al. (2005) report experiments where subjects complete both a binary-choice version and a single-choice version of the EG task and find that subjects make equivalent choices.

characteristics include Female (equaling 1 for women), Low Income (equal to 1 if the subject falls in the lower third of the income distribution), and Young (equal to one if age is below 25). The variable Low Math Score identifies people with low for mathematical ability and is equal to 1 for scores below one standard deviation below the sample mean score (corresponding to a score of 40 out 100); 16% of the subjects fall into this category. A discrete (rather than continuous) measure was employed for math ability because of our strong prior about a minimum necessary level of math ability to understand the experimental decisions. Similarly, Young and Low Income, are defined as dichotomous variables given our original hypotheses.

To study the time stability of estimates in the two experimental measures, we separate the 156 subjects who participated in both the original experiment as well as in re-test experiment (identical to the original) that took place 6 months later. These subjects were selected ahead of time and invited to participate in the retest. Every subject who was invited turned up for the second experiment, alleviating any concerns about sample selection.

3. Methodology

The statistical procedure for obtaining structural estimates for the utility function of the CRRA form is as follows. Subjects faced with several binary choices between two gambles are assumed to have a utility for money (M) given by $U(M \mid r)$, where r denotes the coefficient of relative risk aversion. For each binary choice between gambles subjects are assumed to conduct an expected utility calculation of the form:

$$EU_i = \sum_k \left[p_k \times U(M_k \mid r) \right], \forall\, k = 1,2$$

for each gamble i, where the probability of occurrence of each amount M_k is denoted by p_k.

Denoting EU_L as the 'left' gamble and EU_R as the 'right' gamble, one can construct a probabilistic choice rule with the ratio:

$$\frac{EU_R^{\frac{1}{\mu}}}{EU_R^{\frac{1}{\mu}} + EU_L^{\frac{1}{\mu}}} \quad (1)$$

The parameter μ allows for deviations from the deterministic choice specified by expected utility theory. As in previous work, μ is interpreted as noise: as $\mu \to \infty$ the choice becomes a random decision whereas as $\mu \to 0$ subjects behave exactly as specified by expected utility theory. In this paper, this parameter plays a key role as part of our analysis is based on the estimated noise parameter of each instrument.

The ratio in (1) forms the basis of a logistic conditional logarithmic likelihood function, denoted as $\ell(r, \mu \mid Y_i)$ that can be maximized with respect to r and μ, where the vector Y_i denotes the actual subjects' choices for either the 'Left" or 'Right' gamble. In order to allow for observed heterogeneity, each of the parameters in the vector $[r\ \mu]'$ is specified as a function of individual characteristics, X_i, with associated coefficient vector β. The resulting modified likelihood function is written as $\ell(r, \mu, \beta \mid Y_i, X_i)$.

For those individuals that were randomly selected to participate in a re-test, we estimate the coefficient of relative risk aversion and its associated noise parameter twice: once for the first experiment (Test) and once for second experiment (Re-Test). Estimation is carried out assuming error clustering within a subject; thus, the reported estimates are robust to correlation across a subject's decisions.

4. Estimation Results

We discuss our results in three parts. First, we present the estimates of the risk and noise parameters and discuss the observed differences across experimental measure, as well as subjects' characteristics (including math ability). The second part discusses in detail the trade-offs that arise when using the two experimental measures considered, with an emphasis on the noise and predictive accuracy of each measure. The third part discusses the time consistency results.

4.1 Risk and Noise Estimates

Table 4 presents the estimates of the coefficient of relative risk aversion as a function of several variables; the noise parameter μ is presented here as a constant. Results indicate that, *ceteris paribus*, the HL measure generates a significantly higher estimate of the risk aversion parameter, approximately 50% larger than that of the EG measure. In this specification, younger persons are less risk averse, and, notably, subjects with low math ability exhibit significantly more risk aversion.

Table 5 displays estimates when the noise parameter is allowed to vary by risk measure, gender, income, age and math ability. The estimates of risk aversion differ from Table 4 when noise is allowed to vary systematically. HL now carries a smaller, though still significant, coefficient in r indicating a higher level of measured risk aversion. HL also shows a highly significant effect on the noise parameter, indicating a higher level of randomness in HL decisions. Women now appear more risk averse, but also have lower noise. Income has no effect in either parameter. Young subjects are no longer less risk averse, but rather exhibit lower noise. Low Math Score continues to have a significant effect on risk aversion, but also has marginally higher noise.

Tables 4 and 5 indicate that there is a significant difference between the estimated risk coefficients for subjects conditional on which type of experimental measure they are engaged in. This is a somewhat unexpected result because the measures are relatively similar, in that both consist of choices among relatively simple gambles and both have very similar payoff ranges. In principle, we would expect that if both measures are accurately measuring the degree of preference and noise heterogeneity, the parameters shown in Table 5 should not vary by experimental measure. It is possible, however, that different experimental measures capture preference and noise heterogeneity differently.

In the field, our experimenters noticed that some subjects appeared to have greater difficulty understanding the task in the HL measure, whereas the one-choice format of the EG measure appeared to be understood more quickly and easily. Thus, we hypothesize that while decisions appear to be noisier with the HL measure, this increased noise is probably attributable to individuals with limited mathematical ability. Adding interaction terms of the HL variable with heterogeneity covariates would allow us to investigate this hypothesis, but we encountered poor convergence in such a model. Alternatively, we computed two separate regressions, one for the HL measure and one for the EG measure. Table 6A contains the results of these regressions.

In the risk parameter, we find that the coefficients are fairly consistent across the two models. The constant is substantially higher in HL, reflecting the previously observed pattern. In both cases, women are more risk averse than men, though this effect is twice as larger in the EG regression. Income and Young are insignificantly related to risk aversion for both measures. Low Math Score carries an insignificant coefficient in the HL regression; this is also true for the EG regression, though the coefficient is quite large and would be marginally significant with a one-tailed test. In the noise parameter, the constant is again much higher in the HL regression, about three times the magnitude of the constant in the EG regression. Gender is weakly significant for EG and strongly significant for HL, indicating lower noise for women. Young carries a negative coefficient in both, but is only significant in the HL regression. Notably, Low Math Score individuals exhibit significantly higher noise in HL but not in EG; moreover the coefficient on this variable is an order of magnitude larger than other coefficients (except for the constant) indicating that limited math ability is an important source of noisy behavior. We next explore further the importance of mathematical ability in the estimation of risk preferences.

4.2 Consistency, Noise and Predictive Accuracy

In this section we investigate three themes. First, we look at the sensitivity of our results to a subset of the population that excludes inconsistent subjects. Then, we take a closer look at the differential noise estimates across instruments, with a focus on low math ability subjects. Finally, we analyze the differential predictive accuracy across instruments, again with an eye on low math ability subjects.

One reason that the estimates for HL and EG in table 6A might differ is because of the way choices are made and coded. Recall that for EG, one gamble is chosen from a set. To model HL and EG decisions in parallel ways, we infer the hypothetical binary choices that result form the chosen gamble in EG. Therefore a subject cannot make inconsistent choices in EG. However, in HL, a subject can make inconsistent choices, by switching between the less and more risky options more than once, or can make "backwards" choices by beginning with B and switching to A choices. In table 6B we identify and remove the 75 inconsistent subjects (8.5% of the sample, including 30 of the 141 low-math subjects), and re-estimate both equations using only consistent subjects.

The pattern of coefficient estimates is very similar to those in Table 6A, with a couple of exceptions. In EG, the slightly higher coefficient on Low Math Score combined with a substantially lower standard error means that it now carries a significant coefficient in r, indicating this group is more risk averse. However, for HL, the coefficient on Low Math Score in r remains insignificant. As expected, for the noise parameter estimates, the coefficient on Low Math Score for HL drops in magnitude and is no longer significant, in line with the higher frequency of inconsistent choices among low-math-score subjects displayed in figure 2. In EG there are no longer any significant determinants of noise: noise does not vary by any of the characteristics. However, in HL, men and older persons continue to have higher noise parameters. Our interpretation of these results is that switching (or making reversed choices) in HL is an indicator that the subject does not understand the task, and so HL generates a less reliable measure of risk aversion for this population. Inconsistent behavior is a much bigger problem for low math ability subjects, as shown in figure 2 above. Critical for purposes of measurement is being able to identify whether low ability individuals are more risk averse, or whether higher risk aversion might be an artifact of the measurement process. This procedure suggests that the correlation may be real. This possible correlation between math ability and risk attitudes is discussed further below.

To visualize the differences in the estimated noise parameters across instruments, we calculate the predicted noise parameter for each subject under both tasks. Figure 3, Panel A presents the cumulative distribution of the predicted noise parameter by experimental measures (HL and EG), and by math ability level (for the HL measure only) based on the estimates in Table 6A. There are striking differences between the two measures' estimated noise. The range of the noise parameter for the EG measure is between 0.025 and 0.059 and for the HL measure between 0.058 and 0.244: there is virtually no overlap in the distributions. For HL, the range of the noise parameter shifts considerably for people with low math ability: 0.160 to 0.244.[8] Panel B of the figure graphs the same relationships using only subjects who made consistent choices in the HL measure based on table 6B estimates. Here the range for EG is between 0.023 and 0.059; for HL it is between .050 and 0.164 for all subjects, and between 0.086 and 0.244 for low ability subjects. Compared to Panel A, the HL distributions are shifted to the left, and are much closer to EG, with a larger overlap in the distributions. Using the distributions in Panel B, one can compute a statistical test of the difference in the *median* noise: at any confidence level, the median noise parameter for low math ability subjects under the HL measure is larger than the median noise parameter under the EG measure.

All else equal, if we interpret the noise parameter as mistakes or (more formally) as deviations from expected utility theory towards random behavior, a risk elicitation method that

[8] For medium and high literacy people, range of predicted noise is closer to that of EG but still with cumulative distributions to the right of that of EG (not shown).

generates a smaller noise parameter should be preferred to any other method. In part because of its "coarseness" (only 5 categories as compared with 10 for HL) the price of less noise in the simpler EG measure may come in the form of a reduced accuracy in its ability to predict subjects' actual choices.

We employ a commonly used measure of *predictive accuracy*: the fraction of choices that are correctly predicted with the estimated parameters. To illustrate, the risky choice (i.e. choosing the right gamble) is predicted when the estimated probability in (1) is greater than 0.5; the less-risky choice (left gamble) is predicted otherwise. Computing the fraction of actual choices that are correctly predicted generates a measure that is bounded between zero and one, with a larger number representing a better fit (or predictive accuracy). Using the estimates from table 6A, the HL measure has a predictive accuracy of 0.84 for all individuals, 0.76 for low math ability subjects and 0.85 for high (i.e. non-low) math ability subjects.[9] The EG measure has a predictive accuracy of 0.72 for all individuals (predictive accuracy does not vary importantly by math ability).

At first, these values would suggest that the HL measure always generates a superior predictive accuracy than the EG measure. However, as pointed out in several occasions, the two measures are different on a few dimensions. Importantly for the predictive accuracy measure, the larger number of choices and the different range of the implied CRRA in the HL task may give it an unfair advantage. To see this, consider the summary statistics presented in table 3. It is clear that very few subjects (43 or 5.33%) choose switch to the risky gamble before decision 5 (i.e. 4 safe choices). One obvious reason for this is the negative implied CRRA for these choices; EG implied CRRA is almost always strictly positive. Also, only very few subjects (4 or 0.5%) choose ten safe choices (not shown in table 3). Intuitively, the almost deterministic behavior in these two ends

[9] These predictive accuracy values are nearly identical when inconsistent subjects are removed (i.e. using table 6B estimates)

makes prediction relatively easier for HL (subjects' choice in the EG measure exhibit is much less deterministic for any given gamble, see table 2).

Our previous point can be seen clearly by observing predictive accuracy for each of the ten decisions in HL (Figures 4a and 4b) and for each of the five (implied) decisions in EG (figures 5a and 5b).[10] In absence of the ideal experimental data (an HL measure with 5 decisions with similar implied CRRA ranges to those of EG) that allows a direct comparison of predictive accuracy between the two measures, we "collapse" the ten HL decisions into five. This can be easily done by excluding 5 decisions from the estimation; this procedure exactly transforms the 10 observed choices into the 5 choices that would have been observed if only these 5 decisions had been shown to subjects. Of course, this means that the implied CRRA ranges are modified into (typically) larger ranges. We carried out several estimations of (1), eliminating a different set of 5 decisions each time; the sets we eliminated tended to include the initial and final decisions in the HL task. Figures 4c and 4d display the predictive accuracy by decision when we exclude decisions 1, 2, 4, 9 and 10 to estimate (1) with HL.[11] The corresponding CRRA ranges for these decisions are much closer to those corresponding to the 5 (implied) decisions in EG, and as a consequence, Figures 4c and 4d are similar to 5a and 5b.

To maximize comparability across the two measures, the following predictive accuracy analysis excludes inconsistent subjects. The more comparable 5-decision HL measure now produces a predictive accuracy value of 0.76 for all subjects and 0.68 for low-math ability subjects.[12] These values appear to suggest that the EG measure may be unambiguously preferred for low-math ability subjects as it generates lower noise and a slightly better predictive accuracy than the HL measure.

[10] Our point is strengthened when inconsistent subjects are excluded (Figure 4b). Also, consistent with figure 3, a larger noise (reduced math ability) is related to a reduced predictive power of the HL estimates.

[11] If, instead, decisions 1, 2, 3, 4 and 10 are excluded, a similar shape is obtained.

[12] Predictive accuracy is inferior if decisions 1-4 and 10 are excluded: 0.74 for all individuals and 0.63 for low-math ability people. The reason for this is that there is a higher predictive accuracy for decision 3 (included in Figures 4c and 4d) than for decision 9. Thus, excluding decisions 1,2, 4, 9 and 10 is a relatively conservative way to collapse HL into 5 decisions.

However, we interpret this result with caution as we can not conduct formal statistical tests on predictive accuracy differences. Despite this limitation, it is interesting to investigate the low math ability cut-off math score that would generate an equal predictive accuracy for both EG-all-subjects and HL-low-math-ability-subjects. For our sample, this occurs at the 25th percentile in the math literacy test (a score of approximately 45%).

Our analysis of the trade-offs between predictive accuracy and noise is clearly less than ideal. However, we believe that this evidence (strengthened by our anecdotal experience in the field) can guide researchers when choosing a risk instrument. For people with high math ability, our analysis suggests that the preferred instrument should be HL as it has better predictive accuracy than EG and there is no significant difference in the median noise. On the other hand, the EG measure appears an unambiguously better instrument (in terms of smaller noise and better fit) for those individuals in our sample who have a low mathematical ability. Of course, the choice of instrument will ultimately depend on whether the researcher's premium is on higher predictive power or less noisy choices.

4.3 Time Consistency of Preferences

Table 7 presents estimates of the CRRA utility specification using data for those individuals that participated twice in the experiment. The dummy variable "Re-Test" takes a value of 1 if the task belonged to the second experiment and zero otherwise. This specification does not include any heterogeneity measures because we encountered convergence problems when these were included.

Results in table 6 indicate that regardless of instrument, the estimates of the CRRA coefficient seem to rise the second time the experiment was taken. However, this difference is not statistically significant for the EG measure, and weakly statistically significant for the HL measure. A similar result is found with the noise parameter where the noise level for the EG measure is not statistically different between the two experiments, whereas it is statistically significant at the 15%

for the HL measure. Importantly, relative to EG, the results suggest a higher degree of parameter instability across time with the HL measure.

5. Conclusion

In this paper we set out to examine the trade-offs that arise when choosing among measurement methods of risk aversion that require different degrees of cognitive of mathematical ability to be properly understood. We make use of a unique field experiment that included a mathematical ability questionnaire to address this central question. We find that a measurement method that is more difficult to understand might be better, *sometimes*. The more complex measurement method developed by HL appears to be better suited for subjects with reasonable math abilities. However, too much comprehensibility (as measured by a smaller noise in decisions) and predictive accuracy may be given up if this complex method is used to estimate risk preferences for subjects with low math ability. Thus, for less able subjects, care must be taken to design experiments that are easily accessed and comprehended. Further, we find some evidence that the simpler task may generate risk preference estimates that are more stable across time. An overall conclusion is that economists should be concerned with the elicitation method when investigating risk preferences, and realize that the 'ruler' matters.

Our results also support the mounting evidence that there is an important degree of preference heterogeneity in the population that needs to be incorporated into how economists think about risk preferences. However, results suggest that different elicitation methods may produce different results on preference heterogeneity. This is a result that has not been discussed so far as it was not part of our original question. However, it is important that we spend some time discussing it as it raises important questions.

While economists' conception of risk attitudes suggests that either experimental measure should produce the same risk measure, differences in cognitive ability may hamper subjects' ability to reveal their true preferences via an experimental task. In fact, we do find that the two tasks yield different risk preference estimates but we are cautious in interpreting this difference. First, we do not know the true underlying risk preferences, so there it is not possible to know which measure is a more accurate depiction of reality. Second, the difference in the estimated risk parameter may be due to other differing aspects of the tasks. While we leave a definitive answer of this question for future research, it is important to note that several researchers have found a relationship between cognitive ability and measured risk aversion. Dohmen, et al 2007 measure risk and time preferences using a representative sample of 1000 German adults, and find a correlation with the score on a widely-used IQ test. Both of these studies focus on IQ. Our point is that a real correlation between ability and risk aversion can be obscured by a task that produces different measures for low-math-skill subjects, as well as the reverse: low math skills can produce an effect that looks like risk aversion.

The results suggest that a simpler, more intuitive measure may provide better accessibility, and so more accurate measures of risk aversion, for subjects with low levels of analytic proficiency. However we think it is important to distinguish between general intelligence and math ability per se. Peters et al examine the specific effect of math skills (as distinct from IQ) and find that more highly skilled subjects are less subject to framing effects. Higher math ability should produce more consistent results across measurement methods. Methods that are designed so that low-math-ability persons can understand and complete them are more likely to allow researchers to find any real correlation between ability and risk attitudes.

Experimental research is ideal for sorting out these issues. One strategy, for example, is to adapt the HL measure to make it simpler and easier for low-math subjects to process. Johnson, et al (2007) report results of an experiment using a new interface that presents the HL choices to subjects in a visual, one-at-a time format. Subjects choose between two circles representing the two gambles, with shaded pie-slice areas representing probabilities and images of stacks of money for payoffs.

(Numbers are also included). They report a substantial increase in consistency. Recent refinements of EG also include visual displays of the gambles and images of money (Eckel et al 2007). When these were used with high school students, we found no relationship between math ability and risk attitudes. These results further illustrate the importance of the task interface for collecting accurate preference information.

References

Andersen, S., G. Harrison, M. Lau, and E. Rutstrom. 2006. "Eliciting Risk and Time Preferences." Working Paper 05-24, Department of Economics, College of Business Administration, University of Central Florida.

Andersen, S., G. Harrison, M. Lau and E. Rutström. (2005). "Lost in State Space: Are Preferences Stable?" *International Economic Review*, forthcoming.

Barr, A. (2003). "Risk Pooling, Commitment, and Information: An experimental test of two fundamental assumptions." Center for the Study of African Economics, Oxford University, WPS2003/05.

Benjamin, D., S. Brown and J. Shapiro (2006), "Who is Behavioral? Cognitive Ability and Anomalous Preferences." Unpublished manuscript, Harvard University.

Berg, J., J. Dickhaut, and K. McCabe (2005), "Risk preference instability across institutions: A dilemma." *Proceedings of the National Academies of Science*, 102, 4209-4214.

Binswanger, H., 1980. "Attitudes toward risk: Experimental measurement in rural India." *American Journal of Agricultural Economics* 62, 395-407.

Binswanger, H., 1981. "Attitudes toward risk: Theoretical implications of an experiment in rural India." *Economic Journal* 91, 867-890.

Dohmen, T., A. Falk, D. Huffman, J. Schupp, U. Sunde, and G. Wagner. 2006. "Individual Risk Attitudes: New Evidence from a Large, Representative, Experimentally-Validated Survey." CEPR Discussion Paper No. 5517. http://ssrn.com/abstract=906866

Eckel, C. 2007. "Experiments on Gender Differences." Forthcoming in *The New Palgrave Dictionary of Economics*, 2nd edition, edited by Steven N. Durlauf and Lawrence E. Blume, Palgrave Macmillan: Basingstoke and New York.

Eckel, C. and P. Grossman. 2002. "Sex Differences and Statistical Stereotyping in Attitudes Toward Financial Risk." *Evolution and Human Behavior* 23(4): 281-295.

Eckel, C. and P. Grossman. 2007. "Forecasting Risk Attitudes: An Experimental Study Using Actual and Forecast Gamble Choices." Forthcoming, *Journal of Economic Behavior and Organization*.

Eckel, C., P. Grossman, C. Johnson, A. Milano, C. Rojas, and R. Wilson. 2007. "Explaining Risk Preferences in High School Students: A Preliminary Look." Presented at the Economic Science Association meetings, Atlanta, June, 2006. Working paper, University of Texas at Dallas.

Eckel, C., C. Johnson, and C. Montmarquette. 2005. "Saving decisions of the working poor: short and long-term horizons." In J. Carpenter, G. Harrison, J. List, eds., *Research in Experimental Economics*, *Volume 10: Field Experiments in Economics,* pp. 219-260.

Eckel, C., C. Johnson, C. Montmarquette, and C. Rojas. 2006. "Debt Aversion and the Demand for Loans for Post-Secondary Education." *Public Finance Review* 35(2): 233-262**.**

Engle-Warnick, J., J. Escobal and S. Laszlo. 2005. "The Effect of an Additional Alternative on Measured Risk Preferences in a Field Experiment." Unpublished manuscript. Experimental 0511003, EconWPA.

Harrison, G., E. Johnson, M. McInnes, and E. Rutström. 2005. "Temporal Stability of Estimates of Risk Aversion." *Applied Financial Economics Letters.* 1: 31-35.

Harrison, G., M. Lau and E. Rutström. 2006. "Dynamic Consistency in Denmark: A Longitudinal Field Experiment." Working paper 05-02, Economics Department, College of Business and Administration, University of Central Florida.

Holt, C., and S. Laury. 2002. "Risk Aversion and Incentive Effects," *American Economic Review*, v. 92 (5): 1644-1655.

Isaac, M. and D. James. 2000. "Just Who Are You Calling Risk Averse?" *Journal of Risk and Uncertainty* 20(2): 177-187.

Johnson, C., C. Montmarquette, and C. Eckel. 2003. "Fostering Adult Education: A Laboratory Experiment on the Efficient Use of Loans, Grants, and Saving Incentives". Social Research and Demonstration Corporation.

Johnson, C., C. Eckel and J. Engle-Warnick. 2007. "An Adaptive Instrument to Elicit Risk Preferences." Working paper, University of Montreal, 2007.

Lowenstein, G., D. Read, and, R. Baumeister. 2003. *Time and Decision: Economic and Psychological Perspectives on Intertemporal Choice.* New York: Russell Sage Foundation.

Peters, E., D. Vastfjall, P. Slovic, C. Mertz, K. Mazzocco, and S. Dicker. 2006. "Numeracy and Decision Making." *Psychological Science*, 17(5): 407-413.

Statistics Canada and OECD. 2003. "Learning a Living: First Results of the Adult Literacy and Life Skills Survey." http://www.statcan.ca/english/freepub/89-603-XIE/2005001/pdf.htm

Tables

Table 1: Sample Summary

Demographic	Urban Sample	Non-Urban Sample
Age 18–24	144	26
Age 25–44	352	88
Age 45–55	160	35
Male	293	57
Female	363	92
Post-Secondary Student	96	5
Unemployed	125	38
Part-time employed	137	33
Full-time employed	219	42
Subtotal	*656*	*149*
High school student sample	80	0
Total	*736*	*149*

Table 2: Eckel-Grossman Gamble Choices
(Subjects choose which gamble to play)

Choice (50/50 Gamble)	Low Payoff	High Payoff	Expected Return	Standard Deviation	Implied CRRA* Range	Fraction of Subjects (%)
Gamble 1	28	28	28	0	3.46 < r	10.7
Gamble 2	24	36	30	6	1.16 < r < 3.46	11.2
Gamble 3	20	44	32	12	0.71 < r < 1.16	39.2
Gamble 4	16	52	34	18	0.50 < r < 0.71	16.8
Gamble 5	12	60	36	24	0 < r < 0.50	11.5
Gamble 6	2	70	36	34	r < 0	10.7

*Coefficient of relative risk aversion.

Table 3: Holt-Laury Gamble Choices

Number of Safe Choices	Implied CRRA Range	Fraction of Choices (%)*	
		All Subjects	Excluding Inconsistent Subjects
0	r < -1.71	1.25	1.36
1	-1.71 < r < -0.95	0.34	0.37
2	-0.95 < r < -0.49	0.45	0.25
3	-0.49 < r < -0.14	3.75	3.35
4	-0.14 < r < 0.15	10.78	10.42
5	0.15 < r < 0.41	12.83	11.41
6	0.41 < r < 0.68	23.16	22.46
7	0.68 < r < 0.97	21.23	22.58
8	0.97 < r < 1.37	11.80	12.03
9-10	1.37 < r	14.41	15.76

*# All Subjects =881, # Inconsistent=75. An inconsistent subject is one that, as he/she moves down the HL decisions, switches back to the safe gamble (A) after having chosen a risky gamble (B).

Table 4: Estimates of risk parameter "r" as a function of characteristics.

Parameter/Variable	Estimate	Std. Err.*	*p*-value
r			
HL	0.2295	0.0353	0.000
Female	0.0081	0.0255	0.750
Low Income (<$30,000)	0.0065	0.0201	0.746
Young (<25)	-0.0695	0.0290	0.017
Low Math Score (<mean – 1 SD)	0.1817	0.0339	0.000
Constant	0.4871	0.0320	0.000
μ	0.0802	0.0070	0.000
Log*L*:		5839.5	
Obs.:		13215	

*Clustered by individual

Table 5: Estimates of risk parameter "r" and noise parameter "μ" as functions of characteristics, pooled data regression

Parameter/Variable	Estimate	Std. Err.*	*p*-value
R			
HL	0.1514	0.0339	0.000
Female	0.1168	0.0360	0.001
Low Income (<$30,000)	-0.0003	0.0293	0.991
Young (<25)	-0.0104	0.0305	0.734
Low Math Score (<mean – 1 SD)	0.1625	0.0405	0.000
Constant	0.4096	0.0369	0.000
μ			
HL	0.0605	0.0065	0.000
Female	-0.0212	0.0075	0.005
Low Income (<$30,000)	0.0024	0.0066	0.717
Young (<25)	-0.0149	0.0062	0.016
Low Math Score (<mean – 1 SD)	0.0290	0.0178	0.103
Constant	0.0571	0.0062	0.000
Log*L*:		5695.3	
Obs.:		13215	

*Clustered by individual

Table 6A: Estimates of risk parameter "r" and noise parameter "μ" as functions of characteristics, separate regressions for EG and HL instruments

Parameter/Variable	EG Instrument			HL Instrument		
	Estimate	Std. Err.*	*p*-value	Estimate	Std. Err.*	*p*-value
r						
Female	0.2584	0.0658	0.000	0.1275	0.0386	0.001
Low Income	0.0202	0.0511	0.693	-0.0271	0.0388	0.485
Young	-0.0190	0.0552	0.731	0.0415	0.0384	0.280
Low Math Score	0.2344	0.1570	0.136	0.0218	0.0921	0.813
Constant	0.3260	0.0382	0.000	0.5479	0.0356	0.000
μ						
Female	-0.0134	0.0080	0.092	-0.0337	0.0128	0.009
Low Income	0.0002	0.0063	0.979	0.0113	0.0117	0.332
Young	-0.0074	0.0069	0.283	-0.0390	0.0115	0.001
Low Math Score	0.0123	0.0227	0.587	0.1023	0.0466	0.028
Constant	0.0462	0.0045	0.000	0.1302	0.0117	0.000
Log*L*:		2490.3			3172.2	
Obs.:		4405			8810	

*Clustered by individual

Table 6B: Estimates of risk parameter "r" and noise parameter "μ" as functions of characteristics, separate regressions for EG and HL instruments, Consistent subjects only*

Parameter/Variable	EG Instrument			HL Instrument		
	Estimate	Std. Err.**	*p*-value	Estimate	Std. Err.**	*p*-value
r						
Female	0.2629	0.0721	0.000	0.1268	0.0388	0.001
Low Income	0.0108	0.0549	0.684	-0.0164	0.0420	0.696
Young	0.0071	0.0580	0.903	0.0480	0.0384	0.211
Low Math Score	0.2549	0.1159	0.028	0.0673	0.0910	0.459
Constant	0.3245	0.0378	0.000	0.5536	0.0379	0.000
μ						
Female	-0.0143	0.0097	0.140	-0.0331	0.0123	0.007
Low Income	0.0011	0.0073	0.880	0.0092	0.0113	0.415
Young	-0.0085	0.0090	0.342	-0.0360	0.0109	0.001
Low Math Score	0.0098	0.0210	0.642	0.0357	0.0370	0.335
Constant	0.0455	0.0051	0.000	0.1192	0.0114	0.000
Log*L*:		2260.5			2636.9	
Obs.:		4030			8060	

* An inconsistent subject is one that, as he/she moves down the decisions on the HL task, switches back to the safe gamble (A) after having chosen a risky gamble (B).
** Clustered by individual

Table 7: Estimates of risk and noise parameters (r, μ) in test and re-test periods, separate regressions for EG and HL instruments

Parameter/Variable	EG Instrument			HL Instrument		
	Estimate	Std. Err.*	*p*-value	Estimate	Std. Err.*	*p*-value
r						
Re-Test	0.0798	0.0904	0.377	0.1177	0.0798	0.141
Constant	0.4320	0.0607	0.000	0.7066	0.0552	0.000
μ						
Re-Test	-0.0032	0.0079	0.687	-0.0437	0.0282	0.122
Constant	0.0521	0.0051	0.000	0.1121	0.0198	0.000
Log*L*:		920.6			1182.8	
Obs.:		1560			3120	

*Clustered by individual

Figures

Figure 1: Risk and Return of Eckel-Grossman Gamble Choices

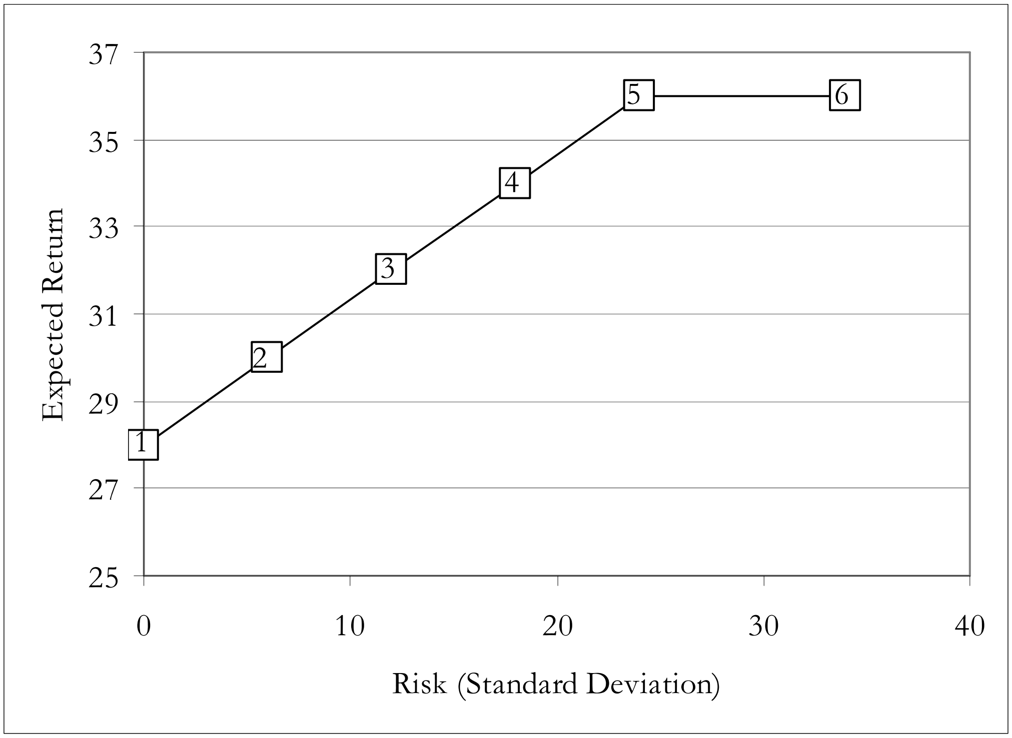
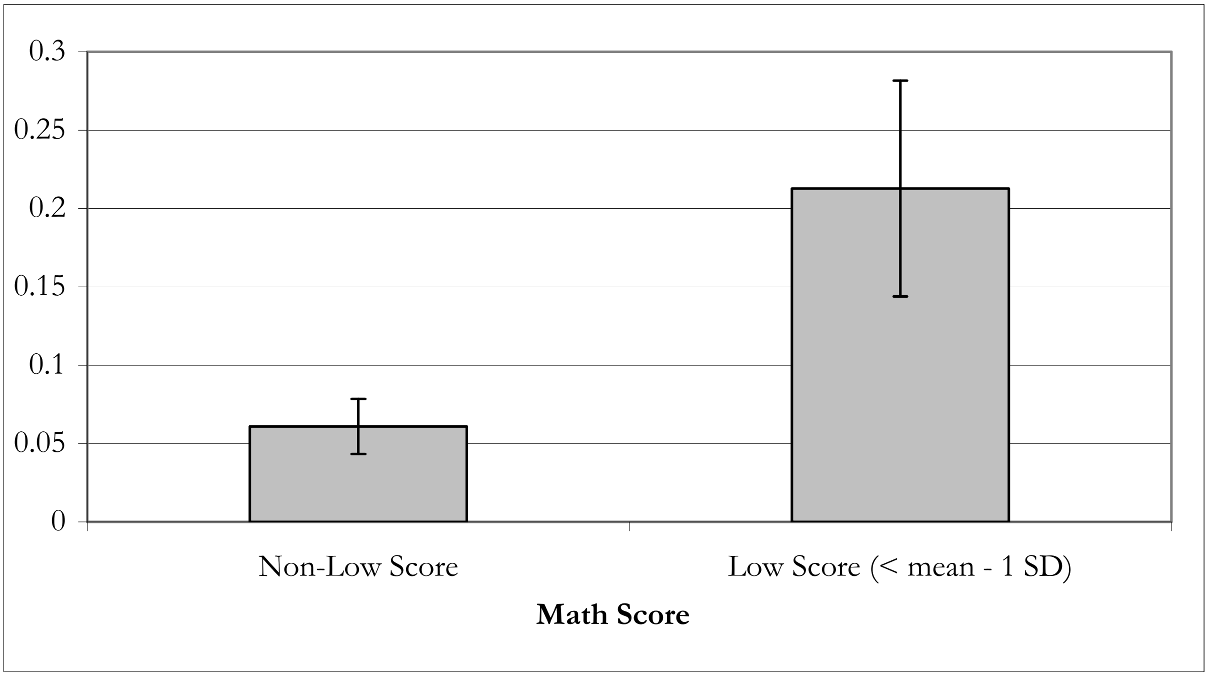


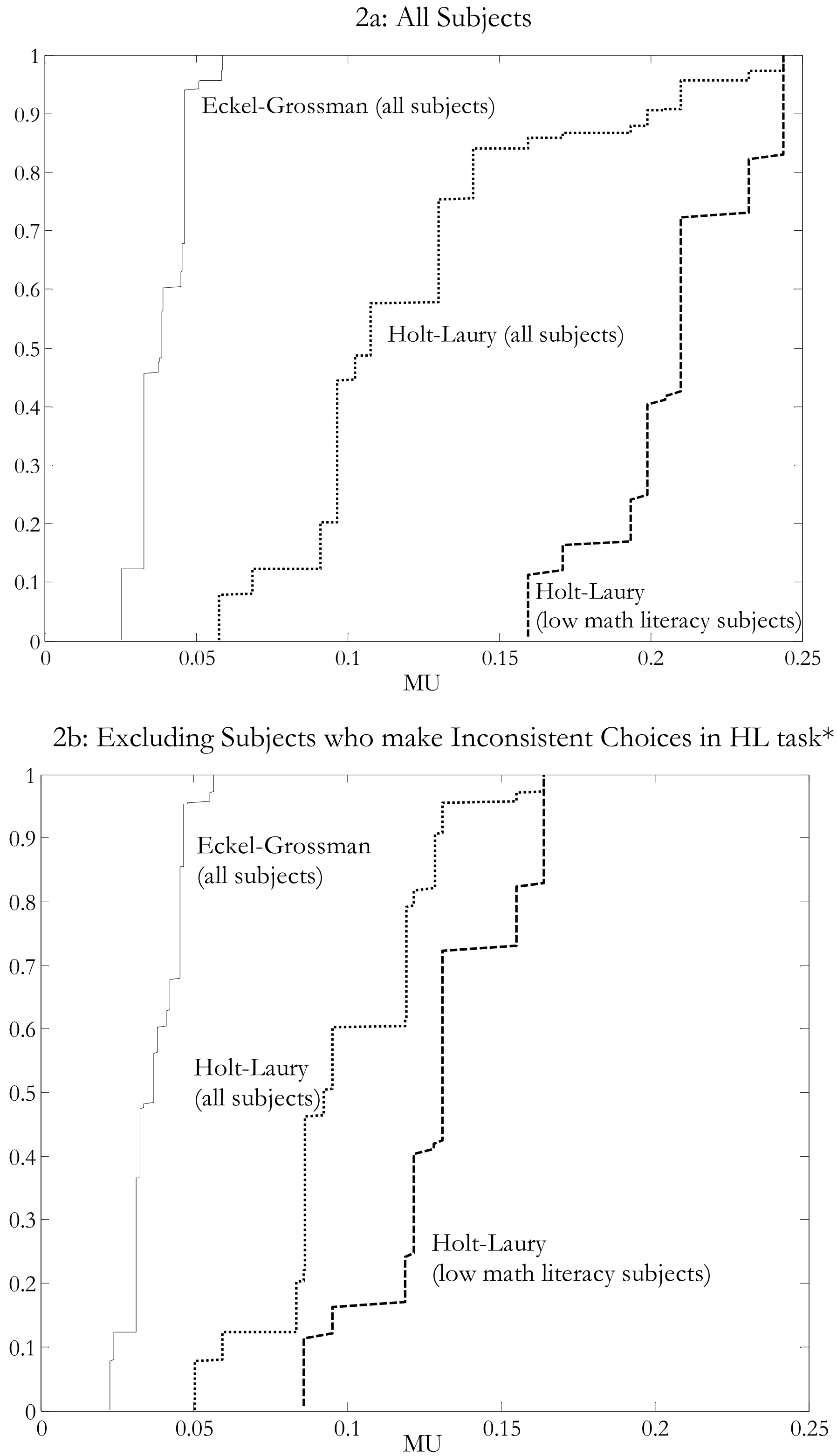
Figure 2: Fraction of Subjects Making Inconsistent Choices in the HL Task*



Note: Vertical bars denote 95% confidence intervals

* An inconsistent subject is one that, as he/she moves down the decisions on the HL task, switches back to the safe gamble (A) after having chosen a risky gamble (B).

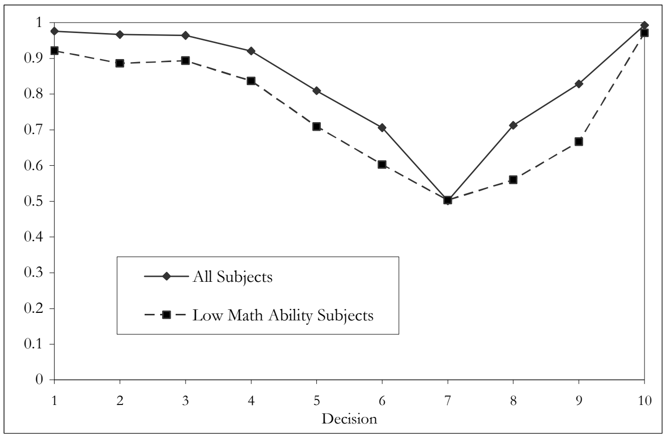
Figure 3: Cumulative Distribution of the Noise parameter by Instrument and by Math Ability



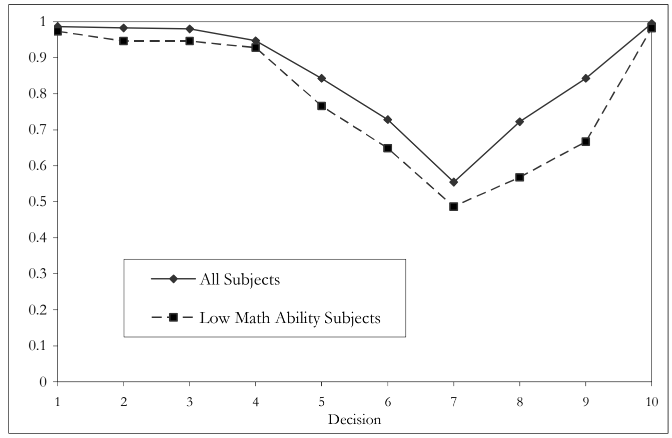
* An inconsistent subject is one that, as he/she moves down the decisions on the HL task, switches back to the safe gamble (A) after having chosen a risky gamble (B).

Figure 4: Fraction of Decisions that are Correctly Predicted in HL Task*

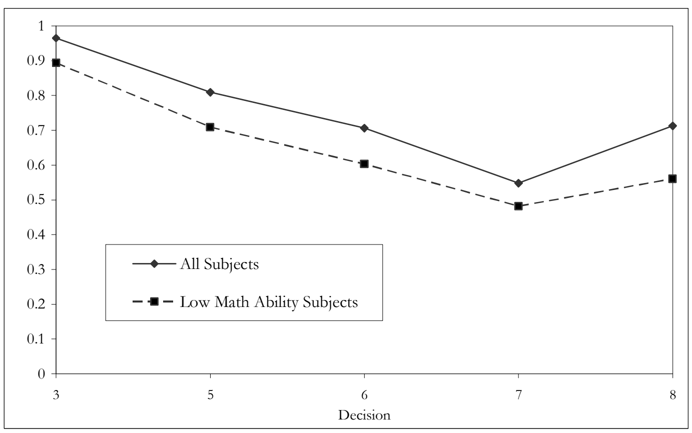
4.a All subjects



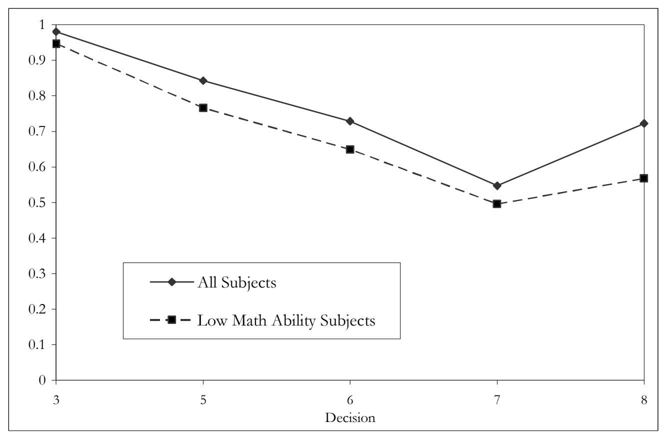
4.b Excluding Inconsistent Subjects



4.c All Subjects: 5 Decisions Only**



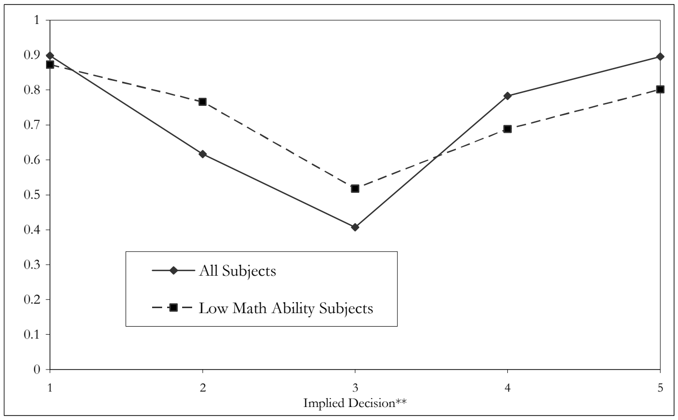
4.d Excluding Inconsistent Subjects: 5 Decisions**



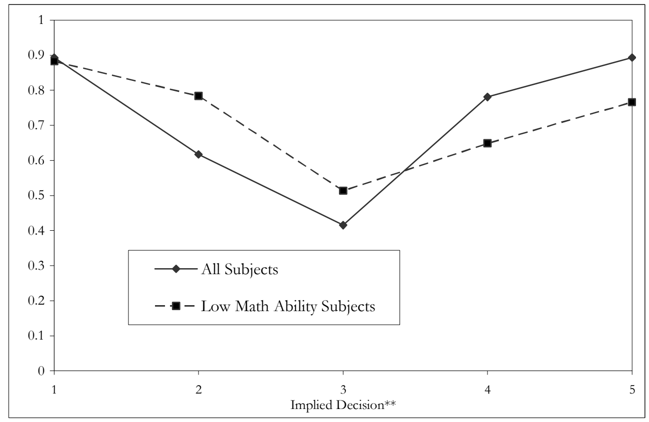
* Using Table 6 estimates
** Excluded Decisions are 1, 2, 4, 9 and 10

Figure 5: Fraction of (implied) Decisions that are Correctly Predicted in EG Task*

5.a All subjects



5.b Excluding Inconsistent Subjects



* Using Table 6 estimates

** See section 2.4 for details on how implied decisions are obtained

Appendix: Decision Forms for Risk Choices

Eckel-Grossman Risk Task

For **Decision 41** you will select from among six different gambles the one gamble you would like to play. The six different gambles are listed below.

- You must select one and only one of these gambles.
- To select a gamble place an **X** in the appropriate box.

Each gamble has two possible outcomes (ROLL LOW or ROLL HIGH) with the indicated probabilities of occurring. Your compensation for this part of the study will be determined by:

- which of the six gambles you select; and
- which of the two possible payoffs occur.

For example, if you select Gamble 4 and ROLL HIGH occurs, you will be paid $52. If ROLL LOW occurs, you will paid $16.

For every gamble, each ROLL has a 50% chance of occurring.

At the end of the study, if **Decision 41** is randomly selected, you will roll a ten-sided die to determine which event will occur. If you roll a 1, 2, 3,4 or 5, ROLL LOW will occur. If you roll a 6, 7, 8, 9 or 0, ROLL HIGH will occur.

Decision 41

Mark your gamble selection with an **X** in the last box across from your preferred gamble.

	ROLL	Payoff	Chances	Your Selection **Mark only one**
Gamble 1	LOW	$28	50%	
	HIGH	$28	50%	
Gamble 2	LOW	$24	50%	
	HIGH	$36	50%	
Gamble 3	LOW	$20	50%	
	HIGH	$44	50%	
Gamble 4	LOW	$16	50%	
	HIGH	$52	50%	
Gamble 5	LOW	$12	50%	
	HIGH	$60	50%	
Gamble 6	LOW	$2	50%	
	HIGH	$70	50%	

In this next set of 10 decisions, you are given a chance to earn a cash prize today. For each decision, you will choose between playing two Gambles, A and B.

Example

Gamble A		Your Choice A or B	Gamble B	
3/10 of $40 (roll: 1,2,3)	7/10 of $32 (roll:4,5,6,7,8,9,0)		3/10 of $77 (roll: 1,2,3)	7/10 of $2 (roll:4,5,6,7,8,9,0)

Each Gamble is composed of two outcomes. Which one occurs depends on the roll of a ten-sided die. For instance, let's look at Gamble A. You have 3 out of 10 chances to win $40 and 7 out of 10 chances to win $32. If you roll a 1, 2 or 3, (3 chances out of 10) then you win $40. If you roll a 4,5,6,7,8,9,0, (7 chances out of 10) then you win $32.

Decisions 42-51:

	OPTION A		Your Choice **A or B**	**OPTION B**	
Decision 1	1/10 of $40	9/10 of $32		1/10 of $77	9/10 of $2
Decision 2	2/10 of $40	8/10 of $32		2/10 of $77	8/10 of $2
Decision 3	3/10 of $40	7/10 of $32		3/10 of $77	7/10 of $2
Decision 4	4/10 of $40	6/10 of $32		4/10 of $77	6/10 of $2
Decision 5	5/10 of $40	5/10 of $32		5/10 of $77	5/10 of $2
Decision 6	6/10 of $40	4/10 of $32		6/10 of $77	4/10 of $2
Decision 7	7/10 of $40	3/10 of $32		7/10 of $77	3/10 of $2
Decision 8	8/10 of $40	2/10 of $32		8/10 of $77	2/10 of $2
Decision 9	9/10 of $40	1/10 of $32		9/10 of $77	1/10 of $2
Decision 10	10/10 of $40	0/10 of $32		10/10 of $77	0/10 of $2