

Internal and External Validity: Comparing Two Simple Risk Elicitation Tasks

UTTEEYO DASGUPTA¹, SUBHA MANI², SMRITI SHARMA³, SAURABH SINGHAL⁴

April 2019

Abstract: Researchers often need to elicit risk attitudes in field experiments or in laboratory experiments. However, choosing a risk elicitation task can be a source of concern, as risk attitudes have been often shown to vary dramatically across tasks. Using a large sample of approximately 2,000 subjects in a within-subjects design, we compare behavior in two commonly used incentivized risk elicitation tasks – the investment game and the ordered lottery choice game – recognized for their simplicity and ease of implementation. We find that risk attitudes elicited from the two tasks show considerable internal consistency, and importantly, have similar predictive validity for behavior in a different task with built-in uncertainty. We also explore determinants of inconsistency between the two risk tasks.

Keywords: Risk preferences, Experiment design, Elicitation methods, India

JEL classification: C91, C81, D81

¹ Wagner College, and Center for International Policy Studies, Fordham University, NY, USA, utteeyodasgupta@gmail.com (corresponding author)

² Department of Economics and Center for International Policy Studies, Fordham University, NY, USA; Population Studies Center, University of Pennsylvania, PA, USA; IZA, Bonn, Germany, smani@fordham.edu

³ Newcastle University, United Kingdom, smriti.sharma@newcastle.ac.uk

⁴ Lancaster University, United Kingdom, s.singhal1@lancaster.ac.uk

We are grateful to two anonymous reviewers, David Budescu, Gary Charness, and Sean Collins for comments and suggestions. We thank United Nations University-World Institute for Development Economics Research (UNU-WIDER) and International Growth Center-India Central for financial support. The funders were not involved in study design or data collection, analysis, and interpretation of data. Neha Agarwal, Riju Bafna, Piyush Bhadani, Japneet Kaur, and Anshul Yadav provided excellent research assistance. We are also grateful to the staff at the various colleges in the University of Delhi for lending their support in conducting the study. The usual disclaimer applies.

1. Introduction

Researchers are frequently in need of simple tasks that can be adopted in lieu of more elaborate elicitation tasks (e.g., Angerer et al., 2015; Charness et al., 2018; Hermann and Musshoff, 2016; Zhang and Ortmann, 2016). There are various incentivized risk elicitation tasks that have been proposed and used in the literature (see Charness et al., 2012 and Holt and Laury, 2014 for recent reviews). However, choosing *a* task among them is not straightforward in light of previous experimental results that often document risk preferences varying dramatically across elicitation tasks, independent of the subject group being WEIRD or non-WEIRD (e.g., Andersen et al., 2008; Berg et al., 2005; Dulleck et al. 2015; Harrison et al., 2005; He et al., 2017; Hey et al., 2009; Isaac and James, 2000; Starmer and Sugden 1993).¹ Further, there are trade-offs involved in choosing between elicitation tasks that present varying levels of difficulty for subjects. While complex tasks, requiring multiple decisions, and involving varying probabilities, may allow for a more precise classification of risk preferences, they may also induce noise due to poor comprehension by subjects (e.g., Charness et al., 2018; Crosetto and Filippin, 2016; Dave et al., 2010).²

For example, Crosetto and Filippin (2016) use a between-subjects design to compare behavior across four incentivized risk tasks: the multiple price list task of Holt and Laury (2002, HL hereafter), the ordered lottery choice game of Eckel and Grossman (2002, 2008, EG hereafter), the investment game of Gneezy and Potters (1997, GP hereafter), and the bomb risk elicitation task of Crosetto and Filippin (2013).³ They find that estimated risk aversion varies greatly across the tasks, and suggest that task-specific features play an important role in explaining the observed differences. Since they use a between-subjects design, they are unable to comment on preference instability at the individual level. Dave et al. (2010) compare the HL and EG tasks, using a sample of Canadian residents, and suggest that the EG task is simpler in representation, easier to comprehend, and generates more stable preference estimates, especially among subjects with lower numeracy. Reynaud and Couture (2012) compare behavior of French farmers in the HL and EG tasks and find that subjects appear more risk averse in the EG task.

¹ Henrich et al. (2010) coined the term WEIRD (Western, Educated, Industrialized, Rich, and Democratic).

² Although the multiple price list task of Holt and Laury (2002) is deemed as the gold standard, inconsistent behavior is often observed in laboratory as well as in artefactual laboratory settings (e.g., Charness and Viceisza, 2016; Cook et al., 2013).

³ The ordered lottery choice task in the EG task is similar to the method proposed by Binswanger (1980, 1981).

Deck et al. (2013) consider whether domain-specific risk attitudes play a role in explaining the within- and between-subject inconsistency in risk preferences. Using four risk tasks (including versions of the EG and the HL tasks), they find that risk attitudes pertaining to gambling and financial domains explain only a small amount of variation in behavior across the risk tasks.

This lack of internal consistency creates a problem, especially since risk elicitation tasks are often used to analyze behavior in other situations. In particular, risk preferences can have important consequences in very different domains which have in-built uncertainty, such as labor market, health, addictive behavior, migration, and self-employment (e.g., Dasgupta et al. 2015; Dawson and Henley, 2015; Dohmen et al., 2011; Hsieh et al., 2017).

In this paper, we assess the internal and external validity of two frequently used incentivized risk tasks – the investment game (GP task) and the ordered lottery choice (EG task). Although the GP and EG tasks pose similar uncertainties in terms of probabilities, they have different representations, and there is little work carefully comparing behavior across these two tasks.⁴ We use risk attitudes elicited from these two tasks in a within-subjects design, and evaluate their internal validity, i.e., intrapersonal consistency of choices in the ordered lottery choice task and the investment game, controlling for individual-specific unobserved heterogeneity. Importantly, we examine external validity, i.e., whether the risk attitudes based on the two elicitation methods provide similar conclusions for predicting behavior under uncertainty in a third game, a competition game, that has in-built uncertainty. We also explore the role of personality traits (such as Big Five and locus of control) as well as socioeconomic characteristics in explaining the inconsistency in risk preferences elicited through the two tasks.

Our results based on a large sample of approximately 2,000 university students in India indicate that risk attitudes elicited from the investment game predict well, albeit not perfectly, attitudes in the ordered lottery choice task. We also find that elicited risk attitudes from either of the tasks predict behavior in the competition task similarly, and there is no significant difference in their predictive validity. Finally, in examining the determinants of inconsistent behavior across the two risk tasks, we find that male subjects are more likely to be consistent in their risk preferences across our chosen tasks. In terms of personality traits, subjects that score higher on

⁴ In fact the investment task is being used increasingly in developing countries with non-WEIRD subjects (e.g., Gong and Yang, 2012; Cameron et al., 2013; Dasgupta et al., 2015; Gangadharan et al., 2016; Dasgupta et al., 2017).

the Big Five emotional stability scale and internal locus of control scales demonstrate less inconsistency in risk attitudes, whereas subjects scoring higher on the Big Five agreeableness scale exhibit more inconsistency in risk attitudes.

2. Experiment

2.1 Choice of Elicitation Tasks

We use a modified version of the original EG task (2002, 2008) similar to the one used in Dave et al. (2010).⁵ Subjects are asked to choose one of six gambles presented in separate rows where each row represents a gamble with equal chances of receiving a high or a low payoff. Columns 1 and 2 in Table 1 list the high and low payoffs for all rows. Gamble 1 is the safe alternative where the high and low payoffs are identical. In moving down from gamble 1 to gamble 5 there is a linear increase in expected returns as well as an increase in the standard deviation of the payoffs; between gambles 5 and 6, there is only an increase in standard deviation, but no increase in expected return. Under expected utility theory (EUT), risk-averse subjects should choose one of the lower-risk, lower-return gambles (i.e., gambles 1-4), whereas risk-neutral subjects should opt for gambles 5 or 6. Further, those opting for gamble 6 (in the presence of gamble 5) can plausibly be characterized as risk-loving.⁶

In the investment or GP task, subjects are asked to divide an allocation of Rs. 150 between a safe asset and a risky investment. If the risky investment is successful (50 percent chance of success), three times the invested amount is paid to the subject along with the amount set aside in the safe option. If the risky investment is unsuccessful, subjects only received the amount set aside in the safe option. Under EUT, a risk-neutral or a risk-loving person should invest the full amount in the GP task.

[Table 1 here]

Given our interest in using intuitive and easily comprehensible incentivized risk elicitation tasks, we purposefully chose the investment game and the ordered lottery choice task

⁵ The original ordered lottery choice task by Eckel and Grossman (2002, 2008) consisted of five gambles.

⁶ While the addition of this sixth gamble helps separate risk averse agents from risk seeking agents, it does not solve the problem observationally, since risk neutral subjects can still be indifferent between the fifth and the sixth gamble.

to minimize task dependent heterogeneity documented in previous work. The two tasks can be categorized under the ‘investment portfolio’ approach (Holt and Laury, 2014) and are similar in many ways.⁷ First, they are both framed as single decision tasks, in contrast to elicitation tasks that require subjects to make multiple decisions. Second, in both games the lottery probabilities are held constant at 0.5, which plausibly allows for better comprehension of the risk since 50 percent may be easier to understand (compared to, for instance, 30 percent or 75 percent). Third, both tasks have the availability of a safe option, and are typically used to separate out degrees of risk aversion. Finally, recent surveys on the perceived complexity of tasks confirm that subjects find the GP and EG tasks to be similar in terms of their complexity (Crosetto and Filippin, 2016).

2.2 Design and Protocol

Our study design consisted of two parts: in the first part, subjects participated in four incentivized tasks, and in the second part, they completed a survey. Each subject participated in all tasks. The four incentivized tasks were a competition game, a distribution game, and the two risk tasks.⁸ Subjects were always presented with the EG task first, and the investment game next. Subjects did not receive feedback within-tasks, or between the tasks. To avoid wealth effects, only one of the tasks was randomly chosen for payments made at the end of a session. After the incentivized tasks, subjects completed a detailed survey that collected information on their family background, academic performance, performance on a Raven’s progression matrices test, and personality traits. We administered standard inventories for Big Five personality traits (this includes openness to experience, conscientiousness, extraversion, agreeableness, and emotional stability) and for locus of control.⁹

⁷ Table B2 and Figure B1 in Appendix B describe the “investment” interpretation inherent in the EG task. We are grateful to a reviewer for suggesting this interpretation.

⁸ Instructions for the tasks are available in Appendix A.

⁹ *Openness to experience* is the tendency to be open to new aesthetic, cultural or intellectual experiences. *Conscientiousness* refers to a tendency to be organized, responsible, and hard working. *Extraversion* relates to an outward orientation rather than being reserved. *Agreeableness* is related to the tendency to act in a cooperative and unselfish manner. *Emotional stability* is predictability and consistency in emotional reactions with absence of rapid mood changes. *Locus of control* ascertains the extent to which individuals believe they can control events affecting them. Those believing that life’s outcomes are due to their own efforts have a higher score on the locus of control (i.e., an internal locus of control), while those believing that outcomes are due to external factors (such as luck) have a lower score (i.e., an external locus of control).

The experiment was conducted among second and third year undergraduate students at the University of Delhi, India. We conducted 60 sessions with approximately 2,000 subjects, resulting in about 34 subjects per session.¹⁰ Each subject participated in only one session lasting for 75 minutes. All subjects received a show-up fee of Rs. 150. The average additional payment from the chosen task was Rs. 230.¹¹

Summary statistics reported in Table 2 indicate that approximately 50 percent of the sample is male, 57 percent belong to households that have high income (at least Rs. 50,000 per month), and 52 percent of the subjects have parents who have a graduate degree or more. Subjects are mostly Hindu (91 percent), and approximately 28 percent belong to disadvantaged groups of Scheduled Castes and Scheduled Tribes (SC and STs) or Other Backward Classes (OBCs) respectively.¹²

[Table 2 here]

3. Results

3.1 Aggregate Behavior in the Two Tasks

In Column 4 in Table 1, we report the distribution of choices in our modified EG task. In our six-row version of the ordered lottery choice task, 80.84 percent of subjects choose gambles 1-4, 9.77 percent choose gamble 5, and the remaining 9.40 percent choose gamble 6. Dave et al. (2010) report that in their sample 77.9 percent choose gambles 1-4, 11.5 percent choose gamble 5, and 10.7 percent choose gamble 6. In general, our subjects appear slightly more risk averse than their Canadian sample, plausibly due to differences in the subject samples and underlying cultural attitudes (India vs. Canada). Similar to our findings, Crosetto and Filippin (2016), using the original five-gamble version of the EG task find that 81.8 percent of their sample choose gambles 1-4. In Table 2, we see that the average subject in our sample chooses gamble 3 (low payoff = Rs. 60, high payoff = Rs. 132), which is close to the mean (2.79) selected gamble in the EG task for Crosetto and Filippin (2016) (note though this is not strictly comparable, as the

¹⁰ Crosetto and Filippin (2016) have approximately 85 subjects participate in the EG and GP tasks.

¹¹ The exchange rate at the time of running these experiments was USD 1 = Rs. 60. The current minimum wage per day for individuals with a high school degree, but not a college degree is Rs. 429, suggesting our risk experiments were sufficiently incentivized.

¹² While age is considered to be an important determinant of risk preferences, by design, there is not much variation in the data due to our exclusive subject sample of second and third year college students.

number of rows are different in the two cases and consequently the presented gamble choices are different in the two studies).

In the GP task, 96.60 percent of the subjects invest less than the full endowment of Rs. 150, and hence appear risk-averse (see Table 2). The average subject invests approx. Rs. 72 in the risky asset, amounting to 47.97 percent of their endowment. Based on the review of existing work, Charness and Viceisza (2016) find that student subjects invest between 44.67 percent and 70.86 percent in the risky asset in the investment game.

Overall, aggregate behavior in the two tasks suggests that a higher proportion of subjects appear to be more risk averse in the GP task compared to our EG task. This observed greater risk aversion in the GP task possibly stems from mental accounting and an aversion towards investing in an environment where there is an apparent chance of losing the invested amount, a conclusion also shared by Crosetto and Filippin (2016).

Further, when we look at behavior separated by gender in the two tasks, females appear significantly more risk-averse than males. In the EG task, 86.48 percent of females and 75.23 percent of males (Mann-Whitney, p -value < 0.001) exhibit risk aversion.¹³ The average row chosen by males is significantly higher than the one chosen by females (3.28 v. 2.80; Mann-Whitney test, p -value < 0.001). Further, the distributions of gambles chosen by males and females are significantly different as seen in Figure 1A (Kolmogorov-Smirnov test, p -value < 0.001). In the GP task, 94.4 percent of males and 98.84 percent of females do not invest the entire endowment (Mann-Whitney test, p -value < 0.001). In terms of share invested in the risky asset, males invest 51.53 percent of their endowment compared to females who invest only 44.4 percent (Mann-Whitney test, p -value < 0.001). Further, Figure 1B shows that distributions of shares invested are significantly different across genders (Kolmogorov-Smirnov test, p -value < 0.001). The significant gender differences found in our experiments are consistent with previous literature (e.g., Charness and Gneezy, 2012; Eckel and Grossman, 2008; Crosetto and Filippin 2016).

[Figures 1A and 1B here]

¹³ Unless noted otherwise, all statistical tests reported in the paper are based on two-tailed tests.

3.2 Internal Validity: Stability in Risk Attitudes

In this section, we use risk attitudes elicited from the EG and GP tasks in a within-subjects design to evaluate their internal validity, i.e., intrapersonal consistency in the two tasks, controlling for individual-specific unobserved heterogeneity. We assume constant relative risk aversion (CRRA) utility function to derive the coefficient of relative risk aversion (r) for the investment task.¹⁴ We regress the risk aversion coefficients from the EG task on those derived from the investment game to determine stability in preferences across tasks.¹⁵ Since the lottery-based EG task elicits ranges of risk aversion rather than point estimates (as seen in Column 3 of Table 1), we rely on interval regression methods, allowing for the dependent variable to have both upper and lower bounds, or be left or right-censored (Anderson and Mellor, 2009; Deck et al., 2013). A coefficient estimate of 1 on the risk aversion coefficient from the GP task would indicate perfect stability/association across two tasks. A coefficient between 0 and 1 would indicate that there is partial stability/association between the two tasks, and a coefficient of 0 would suggest no association between the two tasks.

Table 3 presents the regression results. We find that the risk preferences in both tasks are highly positively correlated, such that a 1-unit increase in the coefficient of relative risk aversion in the GP task increases the coefficient in the EG task by 0.80. Further, we find this marginal effect to be significantly different from 0, but not different from 1 (chi-squared test, p -value = 0.146), indicating a strong association between elicited risk attitudes from the two tasks.

[Table 3 here]

¹⁴ The derivation of the risk aversion coefficient for the two tasks using the CRRA utility function is provided in Appendix B.

¹⁵ Although CRRA provides a highly stylized class of utility functions under expected utility theory (EUT), we adopt it since the large majority of the experimental literature on this topic uses the CRRA model when evaluating and comparing behavior under the EU framework. Consequently, for our purposes of providing a methodological insight, we want to have a similar framework that allows us to make comparisons between our findings and the existing literature. However, more broadly, EU is not the only framework that one can use in studying decision-making under uncertainty, and plausibly, the findings in this paper could vary depending on the underlying model of decision-making under uncertainty.

3.3 External Validity: Predicting Behavior in the Competition Game

We next examine external validity of the two risk elicitation tasks, i.e., if they predict behavior under uncertainty equally well in a third task that elicits willingness to compete. Competitive behavior explains important choices such as participation in labor market training programs, investment, educational outcomes, employment, and the evolution of gender differences in preferences (e.g., Buser et al., 2014; Dasgupta et al., 2015; Niederle and Vesterlund, 2007, 2011; Niederle, 2016) making the determinants of willingness to compete an important area of study.

Since the competition game has an in-built element of risk, to understand the determinants of willingness to compete, one must naturally control for risk preferences. Existing studies looking at the determinants of willingness to compete therefore almost always control for risk preferences. For instance, Balafoutas et al. (2012) use the price list method based on Dohmen et al. (2010, 2011), and Bartling et al. (2009) use the Dohmen et al. (2010) method to elicit risk preferences. Overall, all these papers note that subjects who are more risk averse shy away from competition.

Our interest here is in examining the impact of risk preference using the EG and GP tasks on selecting into the tournament wage scheme. In Table 4, we examine whether the risk aversion coefficients obtained from the two tasks allow us to arrive at similar conclusions when predicting behavior in this third task, the competition game, which is characterized by uncertainty.

In the competition game, subjects are asked to perform an effort-based task (quiz on adding four two-digit numbers), and to choose between a piece-rate payment scheme where payment depends on their own performance, and a tournament wage scheme where payment depends on their own performance relative to that of a randomly selected student who had performed the task earlier. Choosing the tournament scheme, as opposed to the piece-rate payment scheme, involves an element of risk since the payoff in this case depends on relative performance, and could be zero if the subject's performance falls below that of a randomly matched student. Subjects with greater degrees of risk aversion would then prefer the fixed payment scheme to the tournament wage scheme. Consequently, based on the findings from previous literature, we expect the risk aversion coefficients to be negatively related to competitive behavior as measured by the choice of tournament wage scheme (see Bartling et al., 2009).

[Table 4 here]

We estimate OLS regressions where we regress the choice of payment scheme (takes a value 1 if subject chooses the tournament scheme, 0 if piece-rate scheme) on standardized measures of the risk aversion coefficients, obtained from the EG task and the GP task, to make them more readily comparable across regressions (see Columns 1 and 2, Table 4).

In Columns 3 and 4 we also control for a rich set of socioeconomic and demographic characteristics collected in the survey, such as gender, caste, religion, family income, and parental education (see Panel B, Table 2). We include ability, proxied by subjects' performance on a simple two-digit number addition task in the practice round of the competition game, and performance on the Raven's progression matrices test which measures the ability to solve novel problems. We also control for subject's confidence as a dummy that takes a value 1 if the subject believes ex-ante that her performance in the actual task will exceed those of others in the university, 0 otherwise.

Our specifications also include personality traits (such as Big Five and locus of control) as determinants of competitive choices. Our empirical specification is motivated by recent work that shows that personality traits and other socioeconomic characteristics influence competitive behavior (e.g., Bartling et al., 2009; Dasgupta et al., 2019; Gill and Prowse, 2016).

In Column 1 of Table 4, we find that a one standard deviation increase in the risk aversion coefficient from the EG task is associated with a 7 percent decrease in the probability of choosing the tournament wage scheme. We arrive at a similar conclusion by examining the relationship between competition and the standardized risk aversion coefficient from the GP task reported in Column 2, Table 4, though the effect size is now marginally smaller. These results are robust to the inclusion of the full set of controls in Columns 3 and 4. We further find that there is no significant difference in the risk aversion coefficients between Columns 3 and 4 (z-statistic, p -value = 0.16), indicating that risk aversion, as assessed from both these tasks, performs equally well in predicting competitive behavior.¹⁶ Lastly, the R-squared obtained from these regressions indicated in the bottom panel of Table 4 show that the risk aversion coefficients from the EG task reported in Columns 1 and 3 do only slightly better than the risk aversion

¹⁶ These results are robust to using a probit model.

coefficients from the GP task reported in Columns 2 and 4 in explaining the variation in willingness to compete.

Our results are in contrast to Lönnqvist et al. (2015), who used risk attitudes from the HL task and the non-incentivized questionnaire of Dohmen et al. (2011) to predict behavior in a trust game, and found significant differences in the ability of these risk measures to predict trusting behavior. Coppola (2014), however, uses similar multi-item questions on willingness to take risk as well as their own designed lottery tasks to find that measured risk attitudes have reasonable predictive validity for a multitude of life attitude questions related to health, career, sports, etc.

3.4 Sources of Inconsistency

We next examine inconsistency in risk attitudes using the absolute value of the difference in risk aversion coefficients from the EG task and the GP task.¹⁷ For the ordered lottery choice task (EG task), we use the midpoints of the intervals as in Deck et al. (2008) and Deck et al. (2013). Higher absolute difference in risk aversion coefficients denote greater inconsistency.

We find that males are significantly less inconsistent than females (Mann-Whitney test, $p\text{-value} < 0.001$). In Figure 2, we further present the distribution of the inconsistency variable separately for males and females. We reject the null that the two distributions are identical (Kolmogorov-Smirnov test, $p\text{-value} < 0.001$).

To explore the sources of inconsistency, we regress this measure of inconsistency on gender, family background characteristics, and cognitive ability, the set of controls commonly used to explain variation in risk preferences (Dohmen et al., 2010, 2011; Dasgupta et al., 2017). We also control for personality traits as evidence from psychology suggests that variation in risk attitudes (and consequently, inconsistencies in revealed risk attitudes) can be attributed to differences in personality traits (Deck et al., 2008; Weber and Johnson, 2008; Fréchette et al., 2017; Rustichini et al., 2016).

OLS results reported in Table 5 show that males appear to be relatively less inconsistent across risk elicitation tasks. We find that subjects scoring higher on the Big Five emotional stability scale and the locus of control scale are less likely to exhibit inconsistency in risk

¹⁷ We use absolute values of the raw difference in the risk aversion coefficients so that we can treat both positive and negative deviations similarly. Note that the raw difference in risk aversion coefficient has no greater predictive power over the absolute difference in the risk aversion coefficients.

preferences across tasks. We also note that subjects scoring higher on the Big Five agreeableness scale exhibit more inconsistency in risk preferences across tasks. Parental education, family income, caste, and religion have no power in explaining the variation in inconsistency between the two tasks in our sample.¹⁸

[Table 5 here]

4. Discussion and Conclusion

Experimental evaluations of welfare policies have taken off in a big way in developing countries, and new evidence suggests that being aware of recipients' risk attitudes is critical for successful policy implementation (Harrison, 2011; Dasgupta et al., 2015). However, choosing a simple task to elicit risk attitudes is particularly challenging in light of evidence from previous studies that shows elicited risk preferences to be highly sensitive to the method used.

We select two of the simpler and yet commonly used risk elicitation tasks that share similarities in representation, do not vary in probabilities, and are found to be easier to comprehend in the previous literature (Dave et al., 2010; Crosetto and Filippin, 2016) to provide evidence on their internal and external validity. Using data on approximately 2,000 student subjects, we find that the ordered lottery choice task and the investment task show considerable levels of internal consistency. We also find that elicited risk attitudes from the two tasks have similar predictive validity vis-à-vis a third task that has in-built uncertainty, establishing external validity. Further, our results indicate that the consistency across our two chosen tasks depends on gender and some personality traits while other socioeconomic characteristics have no explanatory power.

Since our risk elicitation tasks were chosen to be structurally similar on multiple characteristics as explained earlier, our results suggest that some of the inconsistencies in elicited risk attitudes in the literature can indeed stem from different representations of the tasks, leading to heterogeneity in risk perceptions emanating from those tasks as suggested by Crosetto and Filippin (2016) and Charness et al. (2018). Our results suggest that for fixed and commonly

¹⁸ These results are robust to using an alternative definition for inconsistency where we inconsistency takes a value 1 if the risk aversion coefficient in the GP task does not fall in the range of coefficients for the chosen row in the EG task, 0 otherwise. Table B3 of Appendix B shows that results using this alternative definition are similar to those in Table 5, except that they are somewhat weaker in statistical significance .

understood probabilities and similar elicitation methods, subjects do in fact display more consistent risk attitudes. Consequently, given our relatively large subject pool, we feel confident that using either of the two tasks to elicit risk attitudes would provide comparable results. Consequently, a researcher can pick either of these two simple methods to elicit behavior under uncertainty when faced with laboratory or field constraints.

References

- Andersen, S., Harrison, G. W., Lau, M. I., and Rutström, E. E. (2008). Lost in state space: Are preferences stable? *International Economic Review*, 49(3), 1091-1112.
- Anderson, L.R., and Mellor, J.M. (2009). Are risk preferences stable? Comparing an experimental measure with a validated survey-based measure. *Journal of Risk and Uncertainty*, 39, 137-160.
- Angerer, S., Lergetporer, P., Glätzle-Rützler, D., and Sutter, M. (2015). How to measure time preferences in children: A comparison of two methods. *Journal of the Economic Science Association*, 1, 158-169.
- Balafoutas, L., Kerschbamer, R., and Sutter, M. (2012). Distributional preferences and competitive behavior. *Journal of Economic Behavior and Organization*, 83, 125–135.
- Bartling, B., Fehr, E., Maréchal, M.A., and Schunk, D. (2009). Egalitarianism and Competitiveness. *American Economic Review: Papers & Proceedings*, 99(2), 93-98.
- Berg, J., Dickhaut, J., and McCabe, K. (2005). Risk preference instability across institutions: A dilemma. *Proceedings of the National Academy of Sciences*, 102(11), 4209-4214.
- Binswanger, H.P. (1980). Attitudes towards risk: Experimental measurement in rural India. *American Journal of Agricultural Economics*, 62(3), 395-407.
- Binswanger, H.P. (1981). Attitudes towards risk: Theoretical implications of an experiment in rural India. *The Economic Journal*, 91(364), 867-890.
- Buser, T., Niederle, M., and Oosterbeek, H. (2014). Gender, competitiveness and career choices. *Quarterly Journal of Economics*, 129(3), 1409–1447.
- Cameron, L., Erkal, N., Gangadharan, L., and Meng, X. (2013). Little emperors: Behavioural impacts of China's one-child policy. *Science*, 339(6122), 953-957.
- Charness, G., Eckel, C. C., Gneezy, U., and Kajackaite, A. (2018). Complexity in risk elicitation may affect the conclusions: A demonstration using gender differences. *Journal of Risk and Uncertainty*, 56, 1–17.
- Charness, G., and Gneezy, U. (2012). Strong evidence for gender differences in risk taking. *Journal of Economic Behavior & Organization*, 83, 50-58.
- Charness, G., and Viceisza, A. (2016). Three risk-elicitation methods in the field: Evidence from rural Senegal. *Review of Behavioral Economics*, 3(2), 145-171.
- Cook, J., Chatterjee, S., Sur, D., and Whittington, D. (2013). Measuring risk aversion among the urban poor in Kolkata, India. *Applied Economics Letters*, 20(1), 1-9.

- Coppola, M. (2014). Eliciting risk-preferences in socio-economic surveys: How do different measures perform? *Journal of Socio-Economics*, 48, 1-10.
- Crosetto, P., and Filippin, A. (2013). The “bomb” risk elicitation task. *Journal of Risk and Uncertainty*, 47(1), 31-65.
- Crosetto, P., and Filippin, A. (2016). A theoretical and experimental appraisal of four risk elicitation methods. *Experimental Economics*, 19(3), 613-641.
- Dasgupta, U., Gangadharan, L., Maitra, P., Mani, S., and Subramanian, S. (2015). Choosing to be trained: Do behavioral traits matter? *Journal of Economic Behavior & Organization*, 110, 145-159.
- Dasgupta, U., Gangadharan, L., Maitra, P., and Mani, S. (2017). Searching for preference stability in a state dependent world. *Journal of Economic Psychology*, 62, 17-32.
- Dasgupta, U., Mani, S., Sharma, S., and Singhal, S. (2019). Can gender differences in distributional preferences explain gender gaps in competition? *Journal of Economic Psychology*, 70, 1-11.
- Dave, C., Eckel, C.C., Johnson, C.A., and Rojas, C. (2010). Eliciting risk preferences: When is simple better? *Journal of Risk and Uncertainty*, 41(3), 219-243.
- Dawson, C. and Henley, A. (2015). Gender, risk and venture creation intentions. *Journal of Small Business Management*, 53, 501-515.
- Deck, C., Lee, J., Reyes, J., and Rosen, C. (2008). Measuring risk attitudes controlling for personality traits. Working paper, University of Arkansas.
- Deck, C., Lee, J., Reyes, J., and Rosen, C. (2013). A failed attempt to explain within subject variation in risk taking behavior using domain specific risk attitudes. *Journal of Economic Behavior & Organization*, 87, 1-24.
- Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *American Economic Review*, 100, 1238–1260.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants and behavioral consequences. *Journal of the European Economic Association*, 9(3), 522-550.
- Dulleck, U., Fooker, J., and Fell, J. (2015). Within-subject intra- and inter-method consistency of two experimental risk attitude elicitation methods. *German Economic Review*, 16(1), 104–121.

- Eckel, C.C., and Grossman, P.J. (2002). Sex differences and statistical stereotyping in attitudes toward financial risk. *Evolution and Human Behavior*, 23(4), 281-295.
- Eckel, C.C., and Grossman, P.J. (2008). Forecasting risk attitudes: An experimental study using actual and forecast gamble choices. *Journal of Economic Behavior & Organization*, 68(1), 1-7.
- Fréchette, G.R., Schotter, A., and Trevino, I. (2017). Personality, information acquisition, and choice under uncertainty: An experimental study. *Economic Inquiry*, 55(3), 1468-1488.
- Gangadharan, L., Jain, T., Maitra, P., and Vecci, J. (2016). Social identity and governance: The behavioral response to female leaders. *European Economic Review*, 90, 302-325.
- Gill, D., and Prowse, V. (2014). Gender differences and dynamics in competition: the role of luck. *Quantitative Economics*, 5, 351-376.
- Gneezy, U., and Potters, J. (1997). An experiment on risk taking and evaluation periods. *Quarterly Journal of Economics*, 112(2), 631-645.
- Gong, B., and Yang, C-L. (2012). Gender differences in risk attitudes: Field experiments on the matrilineal Mosuo and the patriarchal Yi. *Journal of Economic Behavior & Organization*, 83, 59-65.
- Harrison, G. (2011). Experimental methods and the welfare evaluation of policy lotteries. *European Review of Agricultural Economics*, 38(3), 335-360.
- Harrison, G. W., Johnson, E., McInnes, M. M., and Rutström, E. E. (2005). Temporal stability of estimates of risk aversion. *Applied Financial Economics Letters*, 1(1), 31-35.
- He, P., Veronesi, M., and Engel, S. (2017). Consistency of risk preference measures: An artefactual field experiment from rural China. *Journal of Development Studies*, forthcoming.
- Henrich, J., Heine, S.J., and Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, 466, 29.
- Hermann, D., and Musshoff, O. (2016). Measuring time preferences: Comparing methods and evaluating the magnitude effect. *Journal of Behavioral and Experimental Economics*, 65, 16-26.
- Hey, J. D., Morone, A., and Schmidt, U. (2009). Noise and bias in eliciting preferences. *Journal of Risk and Uncertainty*, 39(3), 213-235.
- Holt, C. A., and Laury, S. K. (2002). Risk aversion and incentive effects. *The American Economic Review*, 92(5), 1644-1655.

- Holt, C. A., and Laury, S. K. (2014). Assessment and estimation of risk preferences, in M. Machina and K. Viscusi (eds.), *The Handbook of the Economics of Risk and Uncertainty*. Elsevier.
- Hsieh, C., Parker, S.C., and van Praag, C.M. (2017). Risk, Balanced Skills and Entrepreneurship. *Small Business Economics*, 48(2), 287-302.
- Isaac, R., and James, D. (2000). Just who are you calling risk averse? *Journal of Risk and Uncertainty*, 20(2), 177-187.
- Lönnqvist, J. E., Verkasalo, M., Walkowitz, G., and Wichardt, P. C. (2015). Measuring individual risk attitudes in the lab: Task or ask? An empirical comparison. *Journal of Economic Behavior & Organization*, 119, 254-266.
- Niederle, M. (2016). Gender. In John H. Kagel, & Alvin E. Roth (Eds.). *The Handbook of Experimental Economics*, Vol. 2. Princeton University Press.
- Niederle, M., and Vesterlund, L. (2007). Do women shy away from competition? Do men compete to much? *Quarterly Journal of Economics*, 122(3), 1067–1101.
- Niederle, M., and Vesterlund, L. (2011). Gender and competition. *Annual Review of Economics*, 3, 601–63
- Reynaud, A., and Couture, S. (2012). Stability of risk preference measures: results from a field experiment on French farmers. *Theory and Decision*, 73, 203-221.
- Rustichini, A., DeYoung, C., Anderson, J.E., and Burks, S.V. (2016). Towards the integration of personality traits and decision theory in explaining economic behavior: An experimental investigation. *Journal of Behavioral and Experimental Economics*, 64, 122-137.
- Starmer, C., and Sugden, R. (1993). Testing for juxtaposition and event-splitting effects. *Journal of Risk and Uncertainty*, 6(3), 235-254.
- Weber, E. U., and Johnson, E. J. (2008). Decisions under uncertainty: Psychological, economic and neuroeconomic explanations of risk preference. In P. Glimcher, C. Camerer, E. Fehr, and R. Poldrack (Eds.), *Neuroeconomics: Decision making and the Brain*. New York: Elsevier.
- Zhang, L., and Ortmann, A. (2016). Pro-social or anti-social, or both? A within- and between-subjects study of social preferences. *Journal of Behavioral and Experimental Economics*, 62, 23-32.

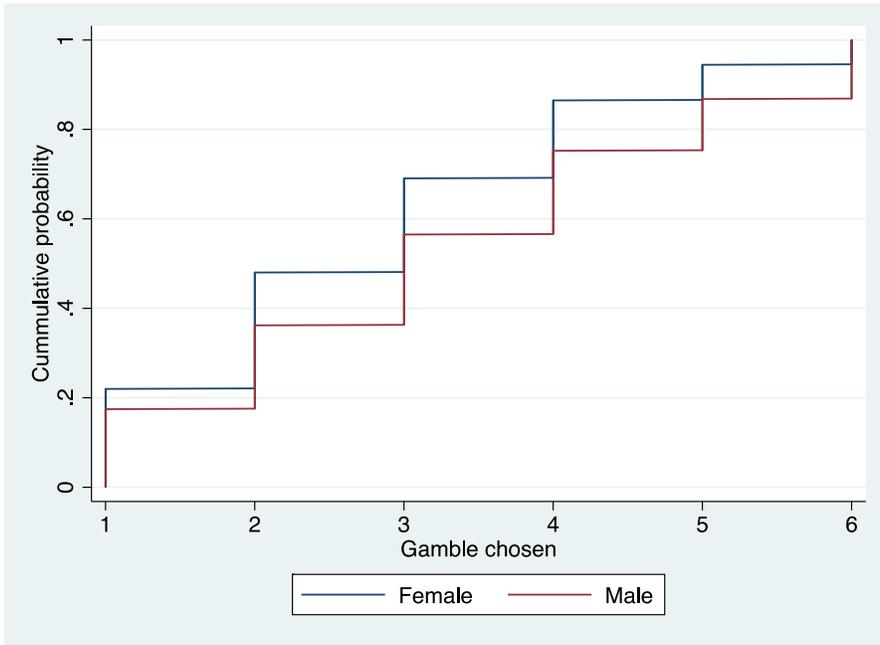


Figure 1A: Gender differences in the gamble chosen in the EG task

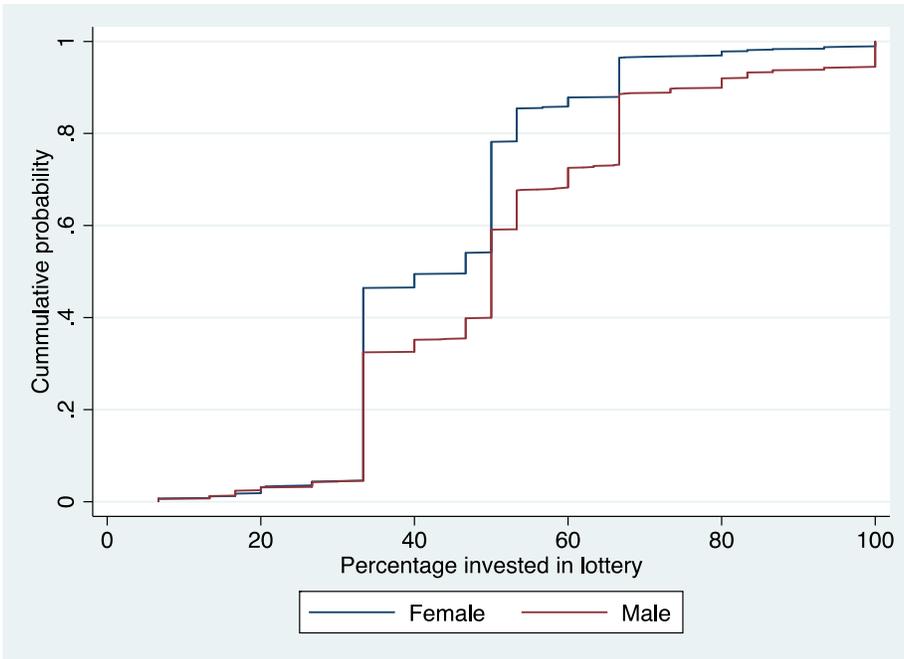
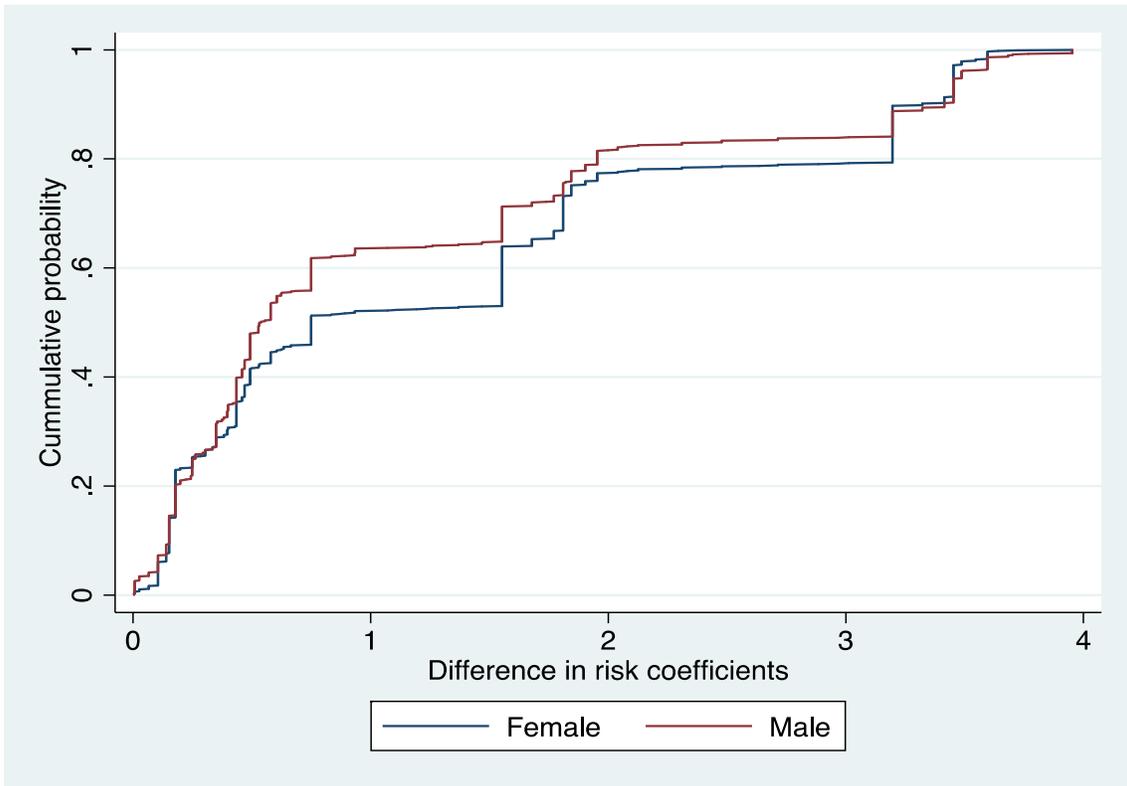


Figure 1B: Gender differences in the share invested in the GP task



Note: The inconsistency measures the absolute value of the difference in risk aversion coefficients elicited from the EG task and the GP task

Figure 2: Gender differences in inconsistency

Table 1: Distribution of Choices in the Modified Eckel and Grossman Task

Choice (50/50 Gamble)	(1) Low payoff	(2) High payoff	(3) Implied CRRA range	(4) All Subjects (%)	(5) Male Subjects (%)	(6) Female Subjects (%)
Gamble 1	84	84	$3.46 < r$	19.74	17.48	22.01
Gamble 2	72	108	$1.16 < r < 3.46$	22.35	18.73	26.00
Gamble 3	60	132	$0.71 < r < 1.16$	20.68	20.29	21.07
Gamble 4	48	156	$0.5 < r < 0.71$	18.07	18.73	17.40
Gamble 5	36	180	$0 < r < 0.5$	9.77	11.55	7.97
Gamble 6	6	210	$r < 0$	9.40	13.22	5.56
Observations				1,915	961	954

Notes: CRRA: coefficient of relative risk aversion

Table 2: Summary Statistics

	Mean (std. dev)
<i>Panel A: Risk measures</i>	
EG task: chosen gamble [1, 6]	3.04 (1.56)
GP task: amount invested in Rs. [0, 150]	71.97 (26.38)
EG task: risk aversion coefficient	1.60 (1.40)
GP task: risk aversion coefficient	0.59 (0.34)
EG task: % risk averse	80.83 (39.36)
GP task: % risk averse	96.60 (18.11)
<i>Panel B: Background characteristics</i>	
Male	0.50 (0.50)
Ability	2.32 (0.89)
Family income	0.57 (0.49)
Parents' education	0.52 (0.49)
Hindu	0.91 (0.28)
Scheduled Caste/Tribe	0.13 (0.33)
Other Backward Classes	0.15 (0.36)
Raven's test score	6.50 (2.27)
Confidence	0.32 (0.46)
Extraversion	4.61 (1.38)
Agreeableness	5.12 (1.15)
Conscientiousness	5.28 (1.26)

Emotional stability	4.55 (1.33)
Openness to experience	5.34 (1.14)
Locus of control	7.28 (1.96)

Observations	1,915
--------------	-------

Notes: The average of the risk coefficient in the EG task is computed over its 6 rows/intervals. The maximum possible scores for the Big Five traits and locus of control are 7 and 13 respectively. The maximum score for Raven's test is 10.

Table 3: Stability of Risk Attitudes

	(1) Relative risk aversion range in the EG task
Risk aversion coefficient in the GP task	0.80*** (0.13)
Constant	1.06*** (0.09)
Null: Risk coefficient in GP task = 1 (p-value)	2.11 (0.14)
Log pseudolikelihood	-3575.38
Observations	1,915

Notes: This table reports coefficient estimates from an interval regression. Standard errors clustered at the session level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: External Validity of Risk Measures

	(1) Compete	(2) Compete	(3) Compete	(4) Compete
Risk aversion coefficient: EG task	-0.070*** (0.011)		-0.056*** (0.012)	
Risk aversion coefficient: GP task		-0.046*** (0.008)		-0.036*** (0.009)
Male			0.143*** (0.024)	0.148*** (0.026)
Ability			0.007 (0.013)	0.006 (0.013)
Confidence			0.089*** (0.022)	0.095*** (0.022)
Family income			0.023 (0.025)	0.024 (0.025)
Parents' education			-0.001 (0.023)	-0.002 (0.023)
Raven's test score			0.001 (0.005)	-0.001 (0.005)
Extraversion			0.005 (0.010)	0.004 (0.010)
Agreeableness			-0.001 (0.010)	-0.004 (0.009)
Conscientiousness			-0.003 (0.013)	-0.005 (0.013)
Emotional stability			0.012 (0.012)	0.015 (0.012)
Openness to experience			0.021** (0.008)	0.020** (0.009)
Locus of control			0.026** (0.012)	0.029** (0.012)
Constant	0.311*** (0.017)	0.311*** (0.017)	0.239*** (0.061)	0.253*** (0.058)
Observations	1,915	1,915	1,915	1,915
R-squared	0.023	0.010	0.067	0.059

Notes: This table reports marginal effects from OLS regressions. Religion and caste dummies also included in this regression. Personality traits are standardized and z-score constructs of these variables used in the regression. Standard errors clustered at session level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Determinants of Inconsistent Behavior

	(1) Absolute difference in risk aversion coefficients
Male	-0.172*** (0.062)
Ability	0.003 (0.033)
Family income	-0.036 (0.071)
Parents' education	-0.046 (0.068)
Raven's test score	0.082 (0.054)
Square of raven's test score	-0.006 (0.005)
Extraversion	-0.020 (0.028)
Agreeableness	0.053* (0.028)
Conscientiousness	0.041 (0.029)
Emotional stability	-0.045* (0.025)
Openness to experience	0.024 (0.026)
Locus of Control	-0.058** (0.024)
Constant	1.047*** (0.212)
Observations	1,915
R-squared	0.018

Notes: This table reports marginal effects from an OLS regression. The dependent variable is the absolute value of the difference in risk aversion coefficients between the EG and GP tasks. Religion and caste dummies also included in this regression. Personality traits are standardized and z-score constructs of these variables used in the regression. Standard errors clustered at session level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Online Appendix: Not for publication

Dasgupta, Mani, Sharma and Singhal

Appendix A: Experiment Instructions

Subject instructions for the Eckel and Grossman Risk Task:

In this task, you will choose one row out of six different rows that are listed on the next page.

Each row consists of a lottery that has two possible outcomes (LOW or HIGH). For every row, HIGH and LOW has an equal chance (50%) of occurring. For example, if you select row 4 and HIGH occurs, you will be paid Rs. 156. If LOW occurs, you will be paid Rs. 48.

Your payment for this task will be determined by:

- Which one of the six rows you select; and
- Which of the two possible outcomes (HIGH or LOW) occur.

If this task is selected for payments, we will roll a six-sided die to determine which outcome will occur. If the die shows 1, 2, or 3 you get the LOW payment from your chosen row. If 4, 5 or 6 appears then you get the HIGH payment from your chosen row.

Refer to the 6 rows at the bottom of this page to answer the following questions before you start:

1. ✓ the row that you think most other students in this room will select?

- Row 1 Row 2 Row 3
 Row 4 Row 5 Row 6

2. If you are selected to receive payments from this task do you think a HIGH outcome will occur for you? (✓ the correct answer)

- a. Yes
b. No

Instructions for completing the task

- You must select one and only one of these six Rows. If you choose more than one Row, your decision will be disqualified if this task is chosen for payment and you will not be paid anything.
- Mark your selected Row with a ✓ in the last box. REMEMBER you choose only ONE of the six rows.

	Outcome	Payoff	Chances	Mark <u>only one</u> Row with ✓
Row 1	LOW	Rs. 84	50%	
	HIGH	Rs. 84	50%	
Row 2	LOW	Rs. 72	50%	
	HIGH	Rs. 108	50%	
Row 3	LOW	Rs. 60	50%	
	HIGH	Rs. 132	50%	
Row 4	LOW	Rs. 48	50%	
	HIGH	Rs. 156	50%	
Row 5	LOW	Rs. 36	50%	
	HIGH	Rs. 180	50%	
Row 6	LOW	Rs. 6	50%	
	HIGH	Rs. 210	50%	

Subject instructions for the Gneezy and Potters Risk Task:

You are given Rs. 150. You have to decide how much of it to keep aside safely, and how much of it to invest in a lottery. There is 50% chance that you win the lottery and 50% chance that you lose. If you win the lottery, you will get 3 times the amount you invest in the lottery; if you lose you will lose the amount you put in the lottery.

Example: Suppose, out of Rs.150 you decide to invest Rs. 100 in the lottery and set aside Rs. 50 safely. If you win the lottery then you will get $3 \times 100 = \text{Rs. } 300$ from the lottery PLUS Rs. 50 that you had set aside safely. If you lose the lottery then you will get Rs. 0 from the lottery PLUS Rs. 50 that you had set aside safely.

If this task is selected for payments, we will roll a six-sided die to determine if you win or lose the lottery. If the die shows 1, 2, or 3, you win the lottery. If it shows 4, 5 or 6, you lose the lottery.

Write your decisions below:

Out of Rs. 150 how much do you want to put in the:

Lottery _____

Keep safely _____

Note: The above two numbers must add up to Rs. 150

Subject instructions for the competition task:

The task is to calculate the sum of four numbers for as many rows as you can in 90 seconds. We will first have a 30 seconds practice round. After that you will get approximately 2 minutes to answer 4 questions asked on the next page. Once everyone has filled his or her responses to the 4 questions, the researcher will announce the start of the actual task for which you will have 90 seconds to complete the task.

You can practice using the following three problems below. The researcher will announce the answers after 30 seconds.

Practice:

Calculate the sum of the four numbers for each of the three rows below. Write your answer in the box with “Total” written in it.

Practice problem 1

35	28	17	42	Total =
----	----	----	----	---------

Practice problem 2

22	45	98	77	Total =
----	----	----	----	---------

Practice problem 3

76	99	56	34	Total =
----	----	----	----	---------

Once you have completed your answers above, please wait till the researcher announces the answers to the questions above.

In the next 2 minutes, please answer the questions on this page. Do not move to the next page until the researcher announces “START”.

1. How many rows do you think you will correctly complete in 90 seconds?

2. As compared to other students in this room, do you think you will be able to complete _____ rows correctly? (✓ the right answer)

- a. More
- b. Less
- c. Same

3. As compared to any other student from Delhi University, do you think you will be able to complete _____ rows correctly? (✓ the right answer)

- a. More
- b. Less
- c. Same

4. For this task, there are two payment schemes. Please select one of them. In case this task is chosen for payments, you will be paid according to your chosen scheme. Place a ✓ for one of the payment options below:

Option 1: you play alone and get Rs. 10 for every correctly solved row.

Option 2: you play against another student from Delhi University (40 students like you have already played this game, and we noted the total number of rows each of them solved correctly in 90 seconds). If this task is chosen for payment, your performance will be matched with the performance of one of the 40 students in the following way: you will be asked to pick a chit. Each chit contains the number of correctly solved rows for one of the 40 Delhi University students. So, there are 40 chits in all. If you correctly solve the same or more rows than the matched student’s performance (the number in the chit), you will get Rs. 20 for every correctly solved answer. If you solve less number of problems than the matched student’s performance then you will get Rs. 0.

Wait till the researcher says “START” to start the task. Stop this task as soon as the researcher announces “STOP”. You will be disqualified from participating further in the game, if you do not stop attempting task 1, after the researcher has announced “STOP”.

TASK:

Calculate the sum of the four numbers for as many rows as you can in 90 seconds. Write your answer in the box with “Total” written in it.

22	98	17	89	Total =
----	----	----	----	---------

44	34	67	66	Total =
----	----	----	----	---------

34	77	67	98	Total =
----	----	----	----	---------

39	87	23	13	Total =
----	----	----	----	---------

99	77	89	45	Total =
----	----	----	----	---------

44	28	80	62	Total =
----	----	----	----	---------

12	25	34	65	Total =
----	----	----	----	---------

67	99	17	66	Total =
----	----	----	----	---------

90	99	78	63	Total =
----	----	----	----	---------

45	67	88	37	Total =
----	----	----	----	---------

78	97	67	25	Total =
----	----	----	----	---------

71	25	37	96	Total =
----	----	----	----	---------

86	78	35	94	Total =
----	----	----	----	---------

26	45	78	29	Total =
----	----	----	----	---------

74	49	37	26	Total =
----	----	----	----	---------

34	42	18	26	Total =
----	----	----	----	---------

86	74	97	33	Total =
----	----	----	----	---------

79	26	45	96	Total =
----	----	----	----	---------

33	67	91	63	Total =
----	----	----	----	---------

43	27	18	95	Total =
----	----	----	----	---------

Appendix B

1. Derivation of the CRRA coefficient from allocation in the GP task

The CRRA coefficients are derived from the investment game in the following way. Assume that individual i divides her budget B between the risky and the safe asset. Assume that L_i is the amount invested in the risky asset. Then the amount invested in the safe asset $S_i = B - L_i$. Assume that I_i is the permanent income of the individual, which can be assumed to be zero without loss of generality. Individual i then maximizes her expected utility, and chooses L_i .

$$\begin{aligned} \text{Max: } E(U_i(L_i, S_i|B)) &= 0.5U_i(S_i) + 0.5U_i(S_i + 3L_i) & (1) \\ &= 0.5U_i(B - L_i) + 0.5U_i(B + 2L_i) \end{aligned}$$

The first order condition for this maximization problem is:

$$\begin{aligned} -0.5U'_i(B - L_i) + 1.5U'_i(B + 2L_i) &= 0 & (2) \\ U'_i(B - L_i) &= 2U'_i(B + 2L_i) \end{aligned}$$

Assuming a CRRA utility function, equation (2) implies

$$(B - L_i)^{-\rho} = 2(B + 2L_i)^{-\rho}$$

Algebraic manipulation yields

$$L_i^* = \left(\frac{2^{1/\rho} - 1}{2^{1/\rho} + 2} \right) B \quad (3)$$

So, the optimal investment in the risky asset (L_i^*) is a function of the coefficient of relative risk aversion (ρ) and the endowment (B). Given L_i^* , we can solve for ρ as

$$\rho = \frac{\ln(2)}{\ln\left(\frac{B + 2L_i^*}{B - L_i^*}\right)} \quad (4)$$

Given B and the optimal investment in the risky asset one can obtain the coefficient of relative risk aversion.

2. Derivation of the CRRA parameters for the Ordered Lottery Choice task (EG task)

We use a modified version of the EG task which adds the sixth gamble as in Dave et al. (2010). The CRRA bounds are found in the task by comparing each gamble (row) to its neighbors and calculating the value of r that generates the same utility level for the payoffs associated with each adjacent gamble. For example, the lower bound of gamble 3 is found by comparing the utility level associated with gamble 4 and gamble 3 (see Table B1 below). Similarly, the upper bound of gamble 3 is found by comparing the utility level associated with gamble 2 with gamble 3. Further a utility function of the form of $U(Y) = Y^{1-r}/1-r$ to represent CRRA is assumed. For our prizes in the modified EG task we increased each original prize from Dave et al. (2010) by a factor of 3. Since we assume CRRA, the risk attitudes of individuals are not affected by a multiplicative transformation of a lottery with a given expected value, by definition of the properties of a CRRA function. This ensures that when we compare the utility levels for each gamble to back out the risk coefficients, we face the original parameters/gamble sizes as in the modified EG task used in Dave et al. (2010). We reproduce them along with our modified payoffs below in Table B1 below. Now, for verification purposes, consider deriving the upper bound for gamble 2 with the prizes used in our experiment. As explained above to get to the upper bound of gamble 2, we compare gamble 1 and gamble 2. Hence, we have,

Expected utility from gamble 1 = expected utility from gamble 2

$$\text{OR} \quad \frac{84^{1-r}}{1-r} = 0.5 \cdot \frac{72^{1-r}}{1-r} + 0.5 \cdot \frac{108^{1-r}}{1-r}$$

$$\text{OR,} \quad \frac{(3 \cdot 28)^{1-r}}{1-r} = 0.5 \cdot \frac{(3 \cdot 24)^{1-r}}{1-r} + 0.5 \cdot \frac{(3 \cdot 36)^{1-r}}{1-r}$$

dividing both sides by 3^{1-r} we get Equation (1):

$$\frac{28^{1-r}}{1-r} = 0.5 \cdot \frac{24^{1-r}}{1-r} + 0.5 \cdot \frac{36^{1-r}}{1-r} \quad (1)$$

Equation (1) is exactly the same comparison we need to calculate r using the Dave et al. (2010) gamble sizes (see Table B1 below) for their gamble 1 and gamble 2. Similar calculations hold for the other rows. Therefore, the CRRA ranges calculated for our paper are the same as those in Dave et al. (2010).

Table B1: Comparison of prizes

Choice (50/50 Gamble)	Size of prizes in our experiment			Size of prizes in the Dave et al. task	
	(1) Low payoff	(2) High payoff	(3) Implied CRRA range	(4) Low payoff	(5) High Payoff
Gamble 1	84	84	$3.46 < r$	28	28
Gamble 2	72	108	$1.16 < r < 3.46$	24	36
Gamble 3	60	132	$0.71 < r < 1.16$	20	44
Gamble 4	48	156	$0.5 < r < 0.71$	16	52
Gamble 5	36	180	$0 < r < 0.5$	12	60
Gamble 6	6	210	$r < 0$	2	70

Table B2: Eckel and Grossman Task with an Investment Game Interpretation

Choice (50/50 Gamble)	(1) Equally likely Payoffs	(2) Amount in Safe Asset	(3) Investment Allocation and Gamble Structure	(4) Implied CRRA range
Gamble 1	84	84	84	$3.46 < r$
Gamble 2	72 or 108	72	$72 + 0.5$ chance of 3×12	$1.16 < r < 3.46$
Gamble 3	60 or 132	60	$60 + 0.5$ chance of 3×24	$0.71 < r < 1.16$
Gamble 4	48 or 156	48	$48 + 0.5$ chance of 3×36	$0.5 < r < 0.71$
Gamble 5	36 or 180	36	$36 + 0.5$ chance of 3×48	$0 < r < 0.5$
Gamble 6	6 or 210	6	$6 + 0.5$ chance of 3×68	$r < 0$

Note: CRRA: coefficient of relative risk aversion

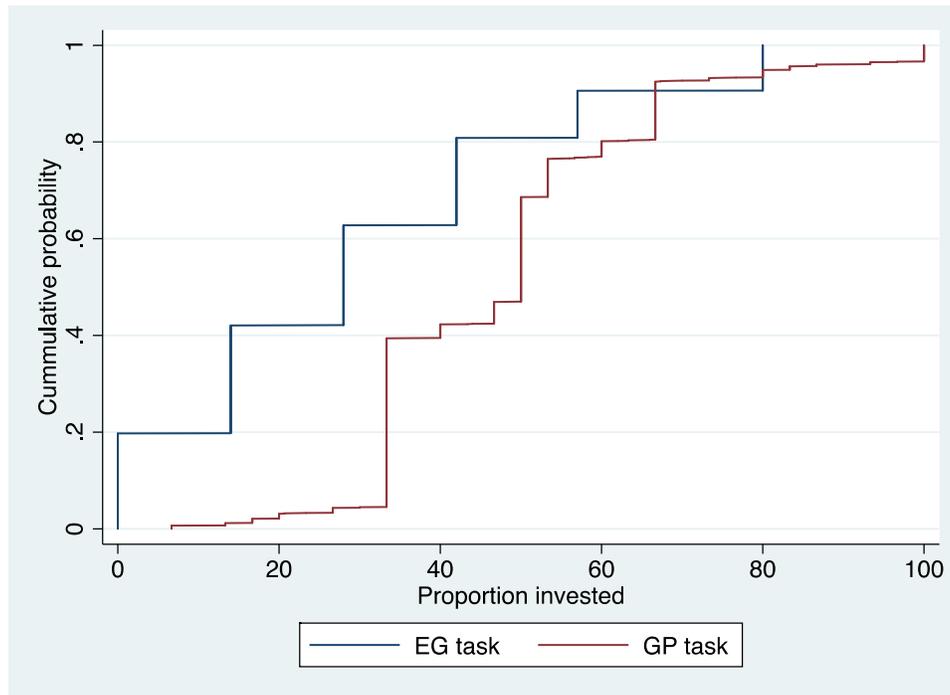


Figure B1: Proportion invested in the lottery in both EG and GP tasks*

*Note: Although the framing of the two tasks appear different, it is possible to construct measures of amount invested in a lottery using the responses from the EG task. For example: A choice of gamble 1 in the EG task can be reinterpreted as investing 0% of your endowment, while a choice of gamble 2 is investing $12/84 =$ about 14% of the endowment.¹⁹

¹⁹ We are grateful to a reviewer for suggesting this interpretation.

Table B3: Determinants of Inconsistent Behavior

	(1) Inconsistency
Male	-0.062*** (0.023)
Ability	0.009 (0.012)
Family income	-0.006 (0.021)
Parents' education	0.002 (0.025)
Raven's test score	0.033 (0.022)
Square of Raven's test score	-0.002 (0.002)
Extraversion	-0.007 (0.010)
Agreeableness	0.004 (0.012)
Conscientiousness	0.003 (0.008)
Emotional stability	-0.015* (0.009)
Openness to experience	0.001 (0.011)
Locus of control	-0.010 (0.008)
Constant	0.690*** (0.080)
Observations	1,915
R-squared	0.014

Notes: This table reports marginal effects from an OLS regression. The dependent variable is a dummy variable that takes a value 1 if the risk aversion coefficient in the GP task does not fall in the range of coefficients for the chosen row in the EG task, 0 otherwise. Religion and caste dummies also included in this regression. Personality traits are standardized and z-score constructs of these variables used in the regression. Standard errors clustered at session level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.