MEASURING RISK AVERSION TO GUIDE POLICY: NATURALISTIC TASKS AND RESPONDENTS

Vinayak V. Dixit, Rami C. Harb,
Jimmy Martínez-Correa,[a] and Elisabet E. Rutström[*]

August, 2013

ABSTRACT When measuring risk attitudes for use in policy analysis it is important to recognize that the context may matter. We propose a method that goes beyond simple context framing using text and pictures and instead relies on Virtual Reality simulations. Such simulations have been shown to improve the accuracy of perceptions in experimental context. We ask whether elicited risk preferences differ when using simulations or stylized lottery tasks, and compare students (a low cost convenient subject pool) to field subjects. While we see initial differences across our 2x2 treatment design, we find that they converge with repetition.

KEYWORDS Policy analysis, risk attitudes, naturalistic tasks, experiments, virtual reality

---

[a] Corresponding author
[*] Authors are listed in Alphabetical order. Senior Lecturer, Research Centre for Integrated Transport Innovation (rCITI), School of Civil and Environmental Engineering, University of New South Wales, Australia, v.dixit@unsw.edu.au, (Dixit); Senior Tolls Consultant, Atkins, NA, Rami.harb@atkinsglobal.com, (Harb); Assistant Professor in Applied Microeconomics, Copenhagen Business School, Department of Economics, Denmark, jima.eco@cbs.dk, (Martínez-Correa); Professor and Director, Dean's Behavioral Economics Laboratory, J. Mack Robinson College of Business and Andrew Young School of Policy Studies, Georgia State University, USA, erustrom@gsu.edu. All online materials available at http://dbel.robinson.gsu.edu. This material is based upon work supported by the Federal Highway Administration under Agreement No. DTFH61-09-H-00012. Any opinions, findings and conclusions or recommendations expressed in this publication are those of the Author(s) and do not necessarily reflect the view of the Federal Highway Administration. We are grateful to Dr. Karen White, Economist in the Office of Transportation Policy Studies at the Federal Highway Administration for guidance and feedback on this research. We are grateful to our other collaborators on the research program including Dr. Glenn W. Harrison, Georgia State University, Dr. Steffen Andersen and Dr. Morten Lau, Copenhagen Business School, and Dr. Essam Radwan, University of Central Florida, for invaluable contributions to the research program. All errors in the paper are those of the authors.

1. INTRODUCTION

Understanding the degree and distribution of risk aversion in a population is essential for policy planning and policy analysis. When measuring such attitudes it is important to recognize that the policy context may matter. We propose a method that goes beyond simple context framing that uses text and pictures and instead relies on Virtual Reality simulations. Such simulations have been shown to improve the accuracy of perceptions in experimental contexts (Fiore, Harrison, Hughes and Rutström [2009]) and therefore can provide a natural bridge between behavior in the laboratory and in the field. We further explore this relationship by comparing behavior of students and a representative sample of the target population.

Risk attitudes enter policy analysis because welfare effects are uncertain and their measurements depend on the risk aversion of people affected by policy changes. For instance, if the policy reduces variations in income, the more risk averse an individual is, the greater is the welfare gain. Also, in some policy areas, such as preventive healthcare, selection into various programs depends on risk attitudes as users perceive the programs as having uncertain effects. Since the effects of the policy depend on these selection processes, predicting and evaluating the effects of the policy depend on the understanding of the distribution of risk attitudes in the population.

There is broad consensus in the behavioral economics and economic psychology literature that human decision makers are generally risk averse. It is also clear that there is a great degree of heterogeneity in these attitudes across individuals (e.g., Andersen, Harrison, Lau, and Rutström [2008] [2010], Gaudecker, Soest, and Wengstrom [2011]). An important question when applying estimates of risk attitudes to policy evaluations is the extent to which these estimates are sensitive to the context. The majority of the evidence has relied on non-contextual tasks such as lottery choices over monetary payoffs (Harrison and Rutström [2008]), or contextually framed hypothetical questions such as job choices with earnings risk (Barsky, Juster, Kimball, and Shapiro [1997]). Barseghyan, Prince, and Teitelbaum [2011] provide evidence from survey data that risk attitudes may not be context neutral. There are, however, few studies to date that directly elicit risk attitudes in contextually framed experiments with real monetary payoffs, and even fewer that do so outside the areas of finance and insurance. We take the notion of context further by observing decisions in contexts where the subject is engaged in a background task that requires both effort and attention and refer to such environments as "naturalistic". One possible reason for this lack of data may be that contextual and naturalistic settings often introduce many confounds that make it difficult to properly identify risk attitudes. The most obvious confound is the absence of control over the probabilities associated with the various outcomes. Exceptions can be found in the area of finance and insurance where it is possible to provide participants with enough naturally occurring information about the probabilities to remove such confounds and still remain in the natural information context (Offerman, Sonnemans, van de Kuilen, and Wakker, [2009], and Wakker, Timmermans, and Machielse [2007]). Virtual reality offers a technology with the same rigorous controls over confounds as other experimental tasks but with naturalistic cues and background tasks that make them particularly relevant for policy settings.

Harrison, Lau and Rutström [2010] illustrate the importance of understanding the degree and distribution of risk aversion when assessing the welfare effects of a policy change where all indirect taxes in Denmark are made uniform. Risk attitudes enter because welfare at the household level is evaluated using Expected Utility, and the more risk averse a household is the smaller is the welfare gain (or the larger is the welfare loss). Anderson and Mellor [2009] and Harrison, Lau and Rutström [2010] report that health seeking behaviors depend on both risk attitudes and the perceived riskiness of the choice options. George, Harrison, Rutstörm and Sen [2012] and Fiore, Harrison, Hughes and Rutström [2009] similarly find that risk management decisions are a function of risk attitudes in experiments using virtual reality renderings of wild fires. These studies, while relying on naturalism in the presentation of the risk, still elicit risk attitudes in a stylized, lottery style task. Thus, only beliefs about the likelihood of fires, and not the attitudes over fire risks, are elicited in a naturalistic setting in these studies. In the area of transportation policy, which is the application used here, an important policy issue is how congestion pricing may provide a solution to the increasing degree of traffic congestion in most metropolitan areas. Responses to increased congestion and to pricing solutions depend on risk attitudes since congestion generates not only longer travel times, but also more unreliable travel times. Dixit, Harrison and Rutström [2013] found little relationship between the distribution of risk attitudes and the propensity for accident risks in naturalistic driving simulator experiments, but they also elicited the risk attitudes using lotteries from students.

This study provides a first attempt at eliciting risk attitudes directly through a controlled but naturalistic task that uses salient monetary payoffs, thus providing evidence that is appropriate for policy analysis. In this study we use a 2 x 2 treatment design in which we compare students and field subjects across traditional stylized experimental tasks, such as binary lottery choices, and naturalistic real-time tasks where the risky choice is embedded in a task that requires both attention and effort. The naturalistic task we use here is driving in a simulated traffic environment. One may expect differences in behavior since these are cognitively quite different tasks. Framing effects are proposed by various dual process models in the psychological literature (Chaiken and Tope [1999], Stanovich and West [2000], Kahneman [2003], Evans and Frankish [2009]). It is possible that the lottery task involves slower, more deliberative cognitive modes and that the simulated driving task involves faster, more affective modes. Mukherjee [2010] offers insights into how dual processes may apply to decisions under risk and uncertainty. He suggests that people vary in their disposition to use either of these cognitive processes, and that task construction can directly affect the weight that either gets in the valuation of a prospect. Further, the choice task in the driving simulator may feel more familiar, at least to participants who have little or no familiarity with the stylized lottery task.

We expect a great deal of heterogeneity in the responses to these two type of tasks, and we therefore ask if our commonly used convenience pool of experimental subjects, viz. university students, behave differently from subjects recruited in the field from the general population. The answer to this question can help us assess if experimenters can conveniently use students to study economic problems and policies where risk attitudes play a key role. Earlier studies based on stylized tasks that compare students to field subjects show that student behavior can be informative of field subject behavior conditional on similarity in demographic compositions and familiarity of the task.[1] Our 2x2 design directly confronts the proposition by List [2007] and Levitt and List [2007] that the representativeness of the environment may be a more important determinant of the validity of observed behavior than the representativeness of the sampled population.

In this study, each experimental task, whether stylized or naturalistic, is a binary choice between two options that differ in risk as defined by objective and known probabilities. The stylized task is a classic Hey-Orme experimental design (Hey and Orme [1994]) with binary choices over abstract monetary lotteries. The naturalistic task is a risky choice over two driving routes that must be made while performing a real-time driving task in the

[1] Cooper, Kagel, Lo and Gu [1999] compare business managers to students in contextually framed tasks and find that managers do better than students only when the task framing is familiar to them. Benz and Meier [2006] report correlations between behavior in lab experiments and actual field situation for charitable giving. Huck and Müller [2012] report that both familiarity with the task stakes and the wider demographic range of the field population matter for comparing students to field subjects in stylized task settings. Andersen, Harrison, Lau and Rutström [2009] report that students predict field subject behavior at least within the demographic ranges represented by the students. Finally, Falk, Meier and Zehnder [2013] found similarities in behavior between students and non-students in the context of social preference games. Thus we see some evidence that student and field subject behavior may differ as a function of familiarity with the environment when the tasks are stylized or based on contextual framing of money choices using text descriptions.

simulator. This experimental design does more than simply present respondents with a contextual frame since the driving task requires effort, attention and real time choices. The two routes differ in the risk of time delays due to congestion and also in consequent money earnings. A subject is tasked with driving from a virtual home to a virtual workplace and in so doing reaches a point where a choice between two routes has to be made. As is the case in the lottery choice task, subjects know the set of outcomes and the probabilities that apply to each. To implement the naturalistic task we opted for a simulator environment, rather than a field driving environment, in order to maximize the control over confounding influences. One important confound in the field is the opportunity cost of the driving time, which we induce in the experiments. The earnings outcomes of the driving choices therefore reflect this induced value of time. Given that the consequences of the route choice in the simulator are probabilistic outcomes that are known to the responder, the simulator task is isomorphic to a standard stylized lottery task. This enables the comparison across tasks.

Our empirical analysis demonstrates that it is important to worry about risk attitudes for policy analysis. We find strong evidence of significant risk aversion, with virtually no risk loving behavior after minimal experience with the tasks. Further, our findings are optimistic about being able to use students to reveal risk attitudes that are relevant for the broader driving population. Our findings are also optimistic about being able to reveal risk attitudes through stylized lottery tasks that are relevant for naturalistic effort based tasks.

2. EXPERIMENTAL DESIGN

Each participant is presented with four lottery tasks and three driving simulator tasks. Each task is a binary choice between a safer and a riskier option, where each option has two possible outcomes. All outcomes are monetary and are fully known to the subject, as are all probabilities. We will first introduce each of these tasks along with a simple decision model for analyzing the responses. We will then explain our recruitment and experimental procedures. Task instructions are available online as Appendix A.[2]

2.2. Binary Lottery Task

Table 1 shows the set of all lottery prizes and probabilities used, and Figure 1 shows an example of what the computer screen looks like in the lottery task. The left lottery is always the relatively safe one, and the right lottery is always the relatively risky one. The same probability of getting the high prize is applied to both lotteries, but it varies across tasks. The lotteries are computerized and each participant first goes through a practice task. This practice task is the same for all participants and it has stakes that are very large to make it clear that it is just practice. No payments are made for the practice task. The program then randomly selects a combination of prizes and probabilities for each paid task from Table 1, thus varying these across participants and tasks. Each person responds to four lottery choices, thus encountering four cases from the set of prizes and probabilities. At the end of each lottery task the participant rolls dice to play out the selected lottery and the earnings are recorded, along with the cumulative earnings.

A subject's choice across the two lottery options in each task can be modeled using a variety of decision models. We illustrate here using the simple Expected Utility Theory (EUT) and assuming a Constant Relative Risk Aversion (CRRA) utility function. We first have the expected utility (EU) of the safe option:

$$EU_S = p\,\frac{{x_L}^{1-r}}{1-r} + (1-p)\,\frac{{x_H}^{1-r}}{1-r} \quad (1)$$

where p is the probability of a low prize, x_L, and therefore (1-p) is the probability of a high prize, x_H, and r is the coefficient of relative risk aversion. The parameter r in the utility function determines both the first- and second-

[2] All online materials are available at http://dbel.robinson.gsu.edu/research-projects/fhwa/

order derivatives. When $r=0$ the utility function is linear, indicating risk neutrality. When $r>0$ the utility function is concave, commonly interpreted as risk aversion, and when $r<0$ the utility function is convex, or risk loving. Similarly we have the expected utility of the risky option:

(2) $$EU_R = p\frac{{y_L}^{1-r}}{1-r} + (1-p)\frac{{y_H}^{1-r}}{1-r}$$

with the same probability, p, of a low prize, y_L, and (1-p) of a high prize, y_H. A participant would then compare these expected utilities and employ some stochastic process for making the choice.

These models can easily be extended to a Rank Dependent Utility (RDU) model by transforming the probabilities into decision weights. [3] First, all outcomes are ranked from best to worst so that the cumulative decision weights are defined in that order. Using a simple power weighting function we then have the following transformation:

(3) $$W_{best} = (1-p)^{\gamma} \text{ and } W_{worst} = 1-(1-p)^{\gamma}$$

where W_{best} is the decision weight given to the best outcome, and W_{worst} is the decision weight given to the worst outcome, which, since there are only two outcomes in our experiments, is simply one minus the decision weight for the best outcome. The transformation function of probabilities into decision weights is assumed to be a power function, for illustrative purposes, but can take on many different shapes. We can then construct the RDU value function for the safe and the risky options, respectively, as follows:

(4) $$RDU_S = (1-p)^{\gamma}\frac{{x_L}^{1-r}}{1-r} + (1-(1-p)^{\gamma})\frac{{x_H}^{1-r}}{1-r}$$

(5) $$RDU_R = (1-p)^{\gamma}\frac{{y_L}^{1-r}}{1-r} + (1-(1-p)^{\gamma})\frac{{y_H}^{1-r}}{1-r}$$

[3] The design does not include both positive and negative incentives to make various route choices, thus making it impossible to separately identify loss aversion.

2.3. Naturalistic Effort Task: Driving Simulator

The driving simulator task is designed to mimic the lottery task in that each subject is presented with a choice between a safer and a riskier option, where both have monetary consequences. The options consist of two routes that each driver can take when driving between home and work: *7th Avenue* which is a route that does not get congested but where a toll is charged, and *9th Avenue* which is a route that can get congested due to the stochastic presence of a school bus but where a toll is not charged. Thus, 7th Avenue is the safe option and 9th Avenue is the risky option. The difference from the lottery task is that there is no congestion risk at all on 7th Avenue. Both of these routes go through the simulator's downtown area.[4]

Subjects are initially introduced to the simulator task via a map of the downtown area, shown in Figure 2. They only see the map perspective during an introduction video.[5] During the choice tasks the subject views everything from the perspective of sitting in the car, which makes the task somewhat familiar to our field subjects who are all commuters in Orlando and Atlanta. The driver's car is initially parked on *B Street* just south of the intersection with *6th Avenue* (labeled "*home*" for illustration). The background task that requires both effort and attention is to drive from this point to the parking lot outside of a warehouse on *F Street* just north of *9th Avenue* (labeled "*work*" for illustration). The drive takes 2 to 4 minutes, depending on which route they take, which scenario they are in, and how they drive. Drivers are requested to follow general traffic rules, such as speed limits, and to assist them in that some additional vehicles have been added to the simulation. These added vehicles also increase the realism of the simulated environment. The choice

[4] The driving simulator is PatrolSim by MPRI, a division of L3 communications. The software is installed on laptop computers (Asus G73JH-A1 and G73AW-A1) under a Windows XP operating system, which is the Windows platform used by PatrolSim. The computers are equipped with a Momo steering wheel and pedal kit for automatic transmission driving.

[5] The introduction video of the driving task that we show to subjects can be found at: http://dbel.robinson.gsu.edu/research-projects/fhwa/

task is binary: they can only take *7th Avenue* or *9th Avenue* between *B Street* and *F Street*. No other options are allowed.

The driver makes the route choice while performing the background task of driving from home to work. It is the binary route choice which reveals the driver's risk attitudes. As the driver reaches the intersection of *7th Avenue* and *B Street* the traffic light always turns red. This is to allow the driver some time to make the choice between turning right to get on *7th Avenue* or to continue straight to take *9th Avenue*. The congestion risk on 9th Avenue is generated by having two simulation scenarios that are selected with some probability. The two scenarios differ only in one aspect: whether a school bus pulls up on *9th Avenue* from *C Street* or not. At the point where drivers make their route choice, the school bus is not visible to the driver, so there is no prior visible information about the presence of a bus. This school bus stops in each of three intersections, causing congestion with delays. In this experiment, the delays are modeled through a monetary, fixed travel time penalty, independent of what the actual time delay was. This makes the cost of delay the same for all participants, conditional on the school bus pulling up. This is an important design aspect which controls for several confounds, such as variable driving skills and perceptions of time use.[6]

Each participant makes three practice drives and 3 paid drives in the simulator.[7] The first practice drive requires them to take *7th Avenue*, and the other two to take the *9th Avenue*, the first time without a school bus and the second time with a school bus. Therefore they have experienced each of the three possible drives. During a break between the practice drives and the paid drives they complete the lottery task.

In the pairwise route choice, *9th Avenue* is the risky option because there is a probability that drivers will encounter a school bus, resulting in a

[6] If the penalty depends on the actual travel time on the routes, then the probability of being paid various amounts would depend on the driving skills and driving habits of the subject, and this probability would no longer be objectively known.
[7] There are additional drives in the simulator following these, but they had a different purpose and are not analyzed here. This is shown in Figure 3.

monetary travel time penalty. This probability varies across subjects and can be 0.3, 0.5 or 0.7. We use decks of cards to implement the random process, where a proportion of the cards have the word "bus" on them. Drivers know this proportion, which makes the probability known to them. The driver selects a card from the deck, and, without showing the subject if the card has the word "bus" on it, the research assistant loads up the relevant scenario. The congestion probability stays the same across all three drives, but varies across subjects.[8]

Similar to the lottery task, probabilities and monetary outcomes are assigned randomly to subjects. The monetary outcomes consist of a "wage", which can be \$4, \$5 or \$6, varied across subjects. If the driver takes *9th Avenue* and if there is no school bus, then this is the payment for that drive; there are no deductions. However, if there is congestion due to a bus the payment drops to \$0.25. The payment under congestion is not related to the actual travel time in this task, but depends only on whether there is a bus or not. This implies that we have full control over factors that otherwise would confound our inferences of risk attitudes: the ability of the driver and any intrinsic value of time. In effect, we are inducing the value of time. If the driver takes *7th Avenue* there is a toll to be paid, implying that the safe option pays less than the maximum payment for the risky option, the wage. The toll is known to the driver before each drive, but varies across the three drives. The toll is selected using three other decks of cards. The design guarantees some variation in these tolls by using different decks of cards for each of the three drives. One deck has a low toll range, one has a medium range, and one a high range. Table 2 shows how the range of tolls varies with wages so that we

[8] This procedure controls for any credibility issue that may arise. The subject chooses a card before each drive but only the experimenter is allowed to immediately see it. The experimenter configures the simulator to run with or without a bus on 9th Ave depending on which card is drawn. After the drive is finished, but only if the subject chose the risky route, the card is revealed and the subject can verify that the simulator displayed the correct scenario with or without a bus. Cards are then reshuffled and the subject proceeds to the next drive.

avoid the possibility that the toll exceeds the wage. We vary the order in which drivers encounter the low, medium and high range.

A subject's choice across the two routes can be modeled using EUT and assuming CRRA utility. We first have the expected utility of the safe option, *7th Avenue*:

$$EU_S = \frac{(w-T)^{1-r}}{1-r} \tag{6}$$

where w is the wage and T is the toll. There is no congestion risk on *7th Avenue* so there is no probability weighting. Similarly, we have the expected utility of the risky option, *9th Avenue*:

$$EU_R = p\,\frac{(0.25)^{1-r}}{1-r} + (1-p)\,\frac{w^{1-r}}{1-r} \tag{7}$$

where w is again the wage, and in the case of congestion the low earnings are always \$0.25, independent of the wage. It is then assumed that agents compare these expected utilities and make a stochastic choice between them. Equations (6) and (7) can be transformed into RDU models in the same way as for the lotteries, by adding probability weighting to the model.

To illustrate an implied pairwise route choice, Table 3 shows an example of the set of 3 choices a driver might face. In this example the probability of encountering a bus is 0.5, the wage is \$5, and the penalty for ending up behind a bus is \$4.75. This implies that the subject is facing a choice of *9th Avenue* that pays \$5 with 50% chance and \$0.25 with 50% chance or a choice of *7th Avenue* that pays \$5 minus the toll for sure. The only value that varies across the three drives is the toll. The last column shows that, for this example, the expected value of taking *7th Avenue* exceeds that of taking *9th Avenue* for the two lower illustrative tolls, so a risk neutral driver would be expected to choose *7th Avenue* for all but the highest of these three toll values (\$3.70). However, a sufficiently risk averse driver will choose to take *7th Avenue* even when the expected value of taking *9th Avenue* exceeds the expected value of taking the other route. In this way we can draw inferences about risk attitudes.

2.4. Recruiting and Experimental Procedures

The tasks analyzed here are part of a larger experiment. In the larger experiment field participants come to four sessions separated by approximately two weeks each. We use data from the first two of the four sessions. During the time between sessions participants complete paid driving tasks on field routes using their own vehicles. Within each session tasks in the simulator are alternated with non-simulator tasks. This is to allow participants to rest and to minimize the chances of simulator induced nausea. The number of tasks, and the order in which they are presented, are partly determined by the need to utilize the participants' time efficiently and by the desire to not cause queuing when accessing the limited number of simulators. Participants arrive individually in a staggered manner and go through the series of tasks sequentially. Figure 3 shows the flow of the tasks in Sessions 1 and 2, the ones that are of relevance here. The tasks that are analyzed here are highlighted in grey in the figure.[9] We can see that two of the lottery tasks and all three driving tasks are conducted during Session 1, and the final two lottery tasks in Session 2. This separation of the lottery tasks across sessions allows us to investigate the effect of earnings accumulation on the estimated risk attitudes.

Instructions are read out individually, and an enumerator is assigned to assist the participant in each task. In the first session participants are first shown a video introducing them to the driving simulator and the routes they may choose from. After that they are able to make three practice drives in the simulator, corresponding to each of the three driving conditions they may encounter in the actual paid tasks. The first non-simulator task, immediately after these practice drives, is a binary lottery task. They are given one practice lottery, followed by two paid lotteries. The next immediate task is the

[9] Instructions for each of the tasks are included in Appendix A.

simulator drive, consisting of 3 separate drives with two possible routes using the same origin and destination, but where the participant may choose different routes each time. Participants then complete several additional tasks in Session 1 that are not of relevance here. Returning to Session 2 they complete other paid tasks before finishing that session with an additional lottery task. Students participate in an experiment consisting of replications of Session 1 and 2 only. These sessions are also separated by two weeks.

Participants are paid for all tasks and earnings are calculated and tracked in a clear and transparent manner. Research assistants helped participants to track of their earnings, making it reasonable to assume that subjects were aware of the cumulative earnings at all times in the experiment. This provides the subjects with important feedback that influences their experience and familiarity with the tasks. We control for any influence that the cumulative earnings have on their choices in the estimations by using an instrumental variable approach. This deviates from the more commonly used payment protocol in risk elicitation experiments where participants make a series of lottery choices without playing them out, and at the end one is randomly selected, then played out and earnings from this one selected lottery is paid.[10] Our solution, to pay for all tasks sequentially, requires that one allows for the possibility that the prizes in a task are mentally integrated with the accumulated earnings when analyzing the data. Simply comparing proportions of choices across the riskier and safer options will not directly reveal risk attitudes since these can be affected by experimental earnings accumulation.

Participants in Orlando, Florida, and Atlanta, Georgia, were recruited by invitation letters. Recipients were randomly selected from the United States Postal Service (USPS) mailing lists, with oversampling from mail carrier routes with median income levels below the state-wide median income

[10] Additionally, Cox, Sadiraj and Schmidt [2011] discuss theoretical problems with employing the traditional random task selection method.

level. Invitation letters directed recipients to our web page where they were instructed to create an anonymous Gmail account to use exclusively for our study to ensure strict privacy.[11] Admission to participate in the study was contingent on respondents being at least 18 years old and holding a valid driver's license and a vehicle with a valid insurance. Participants were informed that driving simulators were to be used in the study and were advised not to participate if they were sensitive to nausea. Four study sites were selected: east Orlando, west Orlando, north-east Atlanta, and north-west Atlanta. The areas were selected because residents there were likely to commute on the study routes used in the larger set of experiments that include a field driving component. The selection of study routes was determined based on how closely substitutable a local road was to an express way. About 140,000 letters were sent out which resulted in 575 participants who showed up to the first meeting. Out of these 283 completed both the lottery and driving simulator tasks analysed here. The remaining subjects were either given a different set of tasks, or dropped out between the first and second session.[12] In addition we recruited students on the campus of Georgia State University during the summer and fall 2012 using the Excen recruiting database. From this effort we obtained complete participation by 205 students.

We collected choice data from 5 cohorts of field subjects: one in each of Atlanta and Orlando during June - July 2011, one in each of Atlanta and Orlando during August - October 2011, and a final one in Atlanta during February – April 2012, for a total of 283 participants. Participants' ages range from 18 years old to 75 years old of which 47% are male and 53% are female. Table 4 summarizes demographic characteristics of both the field subjects and the students. Since all the students are in college, none of them has yet a college degree. 22% of our field subjects have not completed a college degree. The gender distribution is similar across field and student participants, and so

[11] While the Gmail requirement contributed to the low response rate, at the time we started recruitment this was a required privacy layer from the sponsor.
[12] Our retention rate between Session 1 and Session 2 was 91 percent.

is the income distribution, but otherwise there are some differences, as would be expected. The income variable we used for students is that of their parents and we can see that the proportion of students who come from households with an income below the national median ($50,000) is basically the same as for field participants. The age distribution for students is obviously skewed in the younger direction compared to that of field participants with almost all students being younger than 30 years. We also see a larger proportion of African-American participants among the students, which is a function of the high proportion of African-American students on the university campus. Interestingly enough, field participants report similar prior video game experiences as do students.

3. Empirical Analysis

We ask if students are different from field subjects in terms of their risk attitudes as elicited from the lottery task and the driving simulation task. We also ask if behavior in the driving simulator is different from behavior in the lottery tasks as elicited from both field subjects and students. These tasks differ in that the former is a naturalistic, real time task requiring effort, while the latter is a stylized choice task involving no effort beyond that needed for the decision itself. In the models we estimate here, such behavioral differences would be captured by the risk attitudes, i.e. the curvature of the utility functions. We pool all the data across participants and estimate risk attitudes assuming both Expected Utility and Rank Dependent Utility. Before we do, we introduce the data through summary statistics and a simple Probit model of the route choices. Table 5 lists all our exogenous variables.

In all models we include controls for the cumulative earnings throughout the tasks because we are using a payment protocol where earnings are paid for each task. Table 6 shows the earnings accumulated across tasks. We can then test if the utility curvature, as captured by the relative risk aversion, is constant, increasing or decreasing over the domain of earnings in the experiment. Since cumulative earnings are endogenous, we need to use an instrumental variable technique. The endogeneity derives from the fact that the higher the degree of risk aversion, the less money the subject should make on average.[13] We first regress task earnings in each choice task on the exogenous treatment variables for both the safer and the riskier option, particularly on the expected value and the variance of payoffs in each task. In these instrument regressions we also include exogenous variables that reflect how often the bus appeared or the dice selected the high lottery prize in prior tasks as well as the demographic characteristics shown in Table 4 We then estimate our behavioral models on the predicted cumulative earnings from these instrument regressions. The correlations between the predicted earnings and the actual earnings are above .64 for all tasks.

3.1. Propensity to Choose the Risky Option

Table 7 summarizes the proportion of choices that are the risky option by task. Tasks 1-2 are the two lottery tasks in Session 1, tasks 3-5 are the driving simulator tasks in Session 1, and tasks 6-7 are the two lottery tasks in Session 2. All risky choice proportions are between 37% and 51%. The lowest proportion, 37%, is found for the very first task in the experiment, the first lottery, and all subsequent proportions are higher. We also see that the proportion of risky choices is somewhat higher in the driving simulator than in the lottery tasks. Since the lottery tasks and the driving simulator tasks are only similar, not identical, the somewhat higher propensity to choose the risky option in the driving simulator may not reflect a difference in risk attitudes. Differences in payoffs or probabilities will also affect the risky choice propensities, conditional on any risk attitude. We therefore estimate several structural decision models in order to investigate whether risk attitudes differ

[13] In the example in Table 3, for instance, we can see that a sufficiently risk averse individual, who would select the safe option on all three rows, would earn $1.325 less in expected terms than a risk neutral individual.

across students and field subjects, and across lottery and driving simulator tasks.

Table 8 shows four Probit models, one for each combination of type of participant and type of task: Field Participants Lottery, Students Lottery, Field Participants Simulator, and Student Simulator. All coefficients have been transformed to marginal effects. The endogenous variable is the propensity to choose the risky option in the lottery task or *9th Avenue* in the driving simulator task. We also include *t*-tests of the coefficient estimates across models in the last four columns of the table. Due to the many variations in payoffs and probabilities it is not possible to draw inferences about risk attitudes based on the differences in the propensity to take the risky route. Later we will turn to structural decision models where this is not a problem. However, we include the Probit models only as an introduction to those readers who are more familiar with viewing behavior in this way.

The estimates of these models show that behavior in the experiment complies with the basic law of demand, such that when a route becomes more expensive (higher Toll) there are fewer drivers using it. The toll was charged on *7th Avenue*, i.e. the cost of taking the non-congested route. First, we see that the propensity to choose *9th Avenue* in the driving simulator is decreasing in the congestion risk (Prob), and therefore in the expected cost of congestion. It is also increasing in the toll charge on *7th Avenue*, i.e. the cost of taking the non-congested route. We see a similar effect in the lotteries: the propensity to choose the risky lottery is decreasing in the probability of getting the lower prize (Prob). However, field subjects and students differ in how strongly they react to these changes, which may indicate that they differ in risk attitudes. We will investigate this in the models below.

We see a weakly significant increase in the propensity to take the risky option across task repetitions for field subjects in the lottery task and a decrease for students in the simulator (Task). The variable Task is significant in these two cases. In addition, for the field subjects we see a negative task effect in the lotteries done in Session 2, as shown by the negative coefficient on Lott2. Thus there appears to be some learning effects despite the small number of tasks.

The high prize in the risky option has a significant coefficient in each of the four models, however it is positive in lottery tasks and negative in the simulator task. This should not be interpreted as evidence of a behavioral effect, since there are payoff variations across the two tasks. In the simulator task changes in the high risky payoff are due to variations in the wage, and this wage change also affects the payoffs on the safe route. In the lottery task the payoffs in the safe lotteries do not change with the high risky prize, with the exception of the case of $10, which is where all the prizes in both the safe and risky lotteries are increased. Thus, it is not possible to identify if the differences in the coefficients reflect differences in risk attitudes or differences in the payoff parameters for given risk attitudes. This is not a problem for the estimated structural decision models below. However, this is an intentional design aspect allowing us to use lottery tasks that mimic those in the main literature, and driving tasks that frame the payoffs in ways that would be familiar to the participants.

Behavior in the driving simulator is affected by the accumulated earnings, but this is not the case in the lotteries. For the driving simulator the coefficient on CumW is positive and significant. For every dollar increase in CumW, the propensity to take the risky route increases by 3 percentage points for students and by 2 percentage points for field subjects.

To summarize, we see some differences in the Probit models between students and field subjects as well as between the lottery and the driving simulator task that may translate into differences in risk attitudes. To explore this we now turn to the estimates of the structural decision models.

3.2. Structural decision models

Before we discuss our estimates we will introduce the econometric strategy, which is based on Maximum Likelihood estimations of EUT and RDU models, including a behavioral Fechner error. Our approach builds on previous work on structural estimation of risk attitudes by Andersen, Harrison, Lau and Rutström [2008], Harrison and Rutström [2008] and Andersen, Harrison, Lau and Rutström [2010].We explain this methodology in more detail in the online Appendix C.

3.2.1 Econometric Specification

We construct the following latent index that reflects the difference in the subject's valuation of two alternatives in a given binary driving choice (or binary lottery choice). We illustrate here using EUT but this easily extends to RDU. This valuation difference reflects the subject's intensity of preferences over the safer and the riskier options

$$\nabla EU = (EU_R - EU_S)/\mu \tag{8}$$

where, conditional on the type of preference representation (EUT or RDU), EU_i represents the subject's evaluation of the risky or the safe option, as indicated in (8) by the subscripts R or S, respectively. The parameter μ is a structural "noise parameter" used to allow some errors from the perspective of the deterministic model of EUT or RDU We further normalize utility following the contextual utility approach by Wilcox [2011].

The latent index is then linked to observed choices by using a "probit" function that we denote $\Phi(\nabla\text{EU})$:

$$prob\ (choose\ risky\ option) = \Phi(\nabla EU) \tag{9}$$

The index defined by (8) is linked to the observed choices by specifying that the risky option is chosen when $\Phi(\nabla\text{EU})>½$, which is implied by (9). Therefore, this link function models the usual statistical errors. Since the μ parameter scales the difference in valuations either up or down, the value of this parameter informs us of how flat the link function is. By varying the shape of the link function, one can imagine subjects that are more (or less) sensitive to a given difference in the latent index, i.e. the difference in valuations of options. We then maximize the log likelihood conditional on our structural model assumptions.

3.2.2 Expected Utility and Rank Dependent Utility

Table 9 shows our four models estimated using EUT with a CRRA utility function.[14] Table 10 shows the same models under RDU, again with CRRA utility and using the power utility function of equations 6-7. Under EUT, the only parameters to estimate are those of the curvature of utility, since all other variables that influence choice have been controlled through the experimental design. The dominant behavioral mode is risk aversion. Across the four models we see a significantly lower risk aversion for students in the simulator task than we see in any of the other models: this is the only case where the r coefficient is not significantly different from risk neutrality. This is manifested through the utility function being less concave than for students in the lottery task or field subjects in either lottery or driving tasks. However, with task experience students become risk averse in the driving task, as can be seen by the positive coefficient on Task.

But this is not the only difference we see. Even though the other differences in parameter values are not very large, they are significant. Both field and student subjects are more risk averse in the lottery task than in the simulator task, and field subjects are more risk averse than students in both the lottery task and the simulator task. These attitudes converge with experience (shown through the coefficients on Task). The highest initial risk aversion (for field subjects in the lottery task) shows the strongest decrease (-.31), followed

[14] We find a similar story when estimating using a CARA utility function under EUT. Students in driving simulators are initially not risk averse, but with task experience they increase their risk aversion approaching that of the field subjects. We also find the highest degree of risk aversion among field subjects in lotteries. Appendix B shows these estimated models.

by the second highest risk aversion category at a decrease of -.13 (for students in the lottery tasks), with no significant adjustment for field subjects in the simulator and a significant increase for students in the simulator (by +.28). Figure 4 shows the fitted distributions of the estimated risk aversion coefficients for these models. The top left panel compares the risk attitudes of field subjects across lotteries (mean 0.45) and simulators (mean 0.30). There is a left tail for the simulator task that extends into the risk loving range. The top right panel compares the risk attitudes of students across lotteries (mean 0.34) and simulators (mean 0.18). Again, there is a left tail for the simulator task that extends into the risk loving range. The two bottom panels compares students to field subjects in each of the two type of tasks, simulators to the left and lotteries to the right. From the lower left panel it is now clear that the distribution of risk attitudes for students in the simulators have more mass in the risk loving range than does the distribution for field subjects. These tails into the risk loving range disappear as subjects gain experience in the simulator.

The RDU model shows a utility function with similar patterns: we see the least amount of utility concavity for students in the simulator and the most for field subjects in the lottery task, and there is convergence with task experience. The r coefficients are not significantly different across models, but this is because some of the differences are now captured by the γ parameter as well. Even though the curvature parameter on the probability weighting function is not significantly different from 1 for any of the four models, the t-tests indicate that they are significantly different from each other. Students have the lowest γ value and field subjects have the highest γ value. When γ is below 1 the function is concave, indicating probability optimism, and when it is above 1 the function is convex, indicating probability pessimism. Probability optimism and pessimism are aspects of risk attitudes, thus the point estimates of the γ parameter also points to field subjects in the lottery task being the most risk averse and students in the simulator task being the least risk averse.

The pattern of choice propensities in the Probit model shown in Table 8 only partly reveals this pattern of risk attitudes. In the Probit model we see that field subjects are less inclined to choose the risky option compared to students in both the lotteries and the simulators, consistent with the higher risk attitudes we see here. In the Probit model we also see that field subjects and students converge in risk attitudes with experience, also consistent with the models here. The Probit model does not, however, reveal the differences in risk attitudes across the lottery and driving simulator tasks that we see here within each of these subject groups. Students have a significantly higher propensity to choose the risky option but are estimated to reveal a significantly higher degree of risk aversion in the lottery task compared to the driving simulator task. Field subjects have a significantly higher risk aversion in the lottery task than in the simulator task but this is not reflected in the propensity to choose the risky option in the Probit model. This illustrates that if using a Probit model it is necessary for the tasks to have identical incentive structures, but this can be relaxed when estimating structural decision models thus giving the researcher more freedom to design the task to fit the policy context.

Even though field subjects display decreasing risk aversion with experience in the lottery task, there is a discrete increase in risk aversion as they return to the lottery tasks in Session 2 as shown by the positive and significant coefficient on Lott2 in the EUT model. Thus, it appears that field subjects forget their experiences more easily than students, although they once again react to task repetition by lowering risk aversion. We see again, as we did in the Probit model, that there is a significant effect on risk attitudes in the simulator from the earnings accumulation. However, this effect works against convergence, decreasing risk aversion further through a reduction in r in the EUT model. Nevertheless, the increase in risk aversion through task experience is much stronger so the net effect is a convergence in risk attitudes.

The initially much lower risk aversion displayed by students in the simulator may derive from a perception that the simulator is like a video game and that perhaps there is a greater entertainment value in being more willing to take on risk. However, our video game experience variables do not show a significant negative effect on risk attitudes for students. It does, however for field subjects. Field subjects with more video game experience are less risk averse than those with less, but this is only true for those who are younger, since the coefficient on Oldage_3d is positive and significant. It is also important to recognize that students have less field driving experience, and therefore less experience with congestion, than do the field subjects, and that this may affect the initial behavioral difference. The initially much higher risk aversion displayed by field subjects in the lottery tasks is consistent with the proposed Comparative Ignorance heuristic by Fox and Tversky [1995]. They suggest that people behave in a more conservative manner, i.e. are more risk averse, when they feel less knowledgeable about a context or task.

Finally, we look at the Fechner errors, i.e. how sensitive the subjects are to the EU differences between the choice options. The constant term is not significant, implying that the curvature of the likelihood matches the normal density function as defined over the EU differences. There are a few significant covariates for this error term: Students are less sensitive to EU differences in the Session 2 lotteries and field subjects are a little more sensitive to the EU differences as the cumulative earnings increase in the simulator.

Adding demographic variables (listed in Table 4) to the EUT CRRA model we find very few significant effects (see Appendix B). Female students are significantly more risk averse than male students in the driving simulator, implying that our finding of risk neutrality discussed earlier applies only to the male students. However, female students are also more sensitive to differences in EUT, as shown by the positive and significant coefficient on the Fechner error, implying that there is more variability in male than in female student behavior. We find no other significant demographic influence on the estimate of r, but we do find that the income level seems to matter for how sensitive you are to differences in EUT. We find significant income coefficients on μ in both the student lottery model and the field subject simulator model.

4. Conclusions

In this paper we estimate risk attitudes that are relevant for congestion pricing policy and ask two questions that are of importance for measuring risk attitudes that can be used in policy analysis. We are motivated by demonstrations elsewhere in the literature that risk attitudes matter in important ways to welfare estimations. First we ask, how well can students reveal risk attitudes that are representative of the driving population? Students would be expected to be more experienced with video games and less experienced with driving under congestion than the general driving population, and also differ in certain demographic characteristics, such as age and education levels. Do these differences lead to deviations in risk attitudes that are significant? If not, using students to characterize risk attitudes of a broader population may be possible. We also ask if we need to use naturalistic tasks such as driving simulators in order to reveal policy relevant risk attitudes, or can simpler tasks, such as lotteries, reveal the same risk attitudes? Since the simulator tasks have many elements in common with an actual field driving task, such as making route choices in real time and expending effort and attention on the task of driving itself, we take this task as the benchmark. If lottery tasks generate risk attitudes that match those elicited in the driving simulator, perhaps it is possible to forego more expensive and involved task designs in favor of the well-studied, simpler lottery tasks.

Assuming that choices in driving simulators better represent choices in the field we take the risk preferences of field subjects in the simulators as our benchmark. We find significant evidence of risk aversion with an r-coefficient of 0.75. This level of risk aversion implies that assuming risk neutrality would result in serious biases in both welfare assessments and in impact assessments of policy intervention. For the type of stakes we have used in our tasks in this paper, the certainty equivalent of a risky option with an equal chance of a low outcome of $0.25 and a high outcome of $6 would be around $1.67, while the expected value would be $3.13. Thus, employing an assumption of risk neutrality and basing the welfare assessment on the expected value rather than on the certainty equivalent would overestimate welfare by 87%. The implication for impact assessments can be demonstrated with the following numeric example. Assume that the perceived likelihood of a good policy outcome is uniformly distributed in the population over the range [0.1, 0.9]. If the safe status quo outcome is valued at $2 and the risky policy outcome is the one stated above, we calculate the propensity to choose the safe outcome as 1/3 under risk neutrality and .56 when the r-coefficient is 0.75. Thus, an impact assessment of a policy intervention with a risky outcome like this one would overestimate the acceptance rate by 40%. This demonstrates the importance of accurately incorporating risk attitudes and their distribution in the population in policy analysis and impact assessment.

We find that students initially behave very differently from field subjects in the simulators, revealing a much lower aversion to risk and thus a stronger propensity to choose the risky, congested route. However, if students are given some experience, therefore getting feedback from their actions, they will act more consistently with field subjects. Thus, our findings are optimistic about being able to use students in simulators to reveal risk attitudes that are relevant for the broader driving population. Only minimal experience is required for students to express risk attitudes that are similar to those of field subjects, conditional on demographics. While we see that field subjects are initially much more risk averse in the lottery tasks than in the simulator, experience with the lotteries result in a fast convergence towards the risk attitudes revealed in the simulator. Thus, our findings are also optimistic about being able to reveal risk attitudes through stylized lottery tasks that are relevant for naturalistic effort based driving tasks.

It is important to recognize the limitations of these findings. The payoff structure in these tasks was designed to be relevant for congestion pricing. For example, the range of tolls in the driving simulator was $0.50 - $5.50, which is reasonable in comparison to most road tolls found in the field. At first glance, the wages used in the simulator may seem small, $4 - $6, but considering that the time it takes to earn these wages is only about 10 minutes they correspond to hourly wages of around $24 - $36. While these payoffs match the policy questions we focus on here, for other policy questions they may not be ideal. Thus, further research would be necessary in order to establish how general these findings are with respect to a broader set of policy applications. We expect that there is a large set of policies for which these payoff ranges apply. Policies involving changes in sales or value added taxes, or user fees for many public services are examples that would involve such payoffs.

Figure 1: Screen Shot For Lottery Practice Task

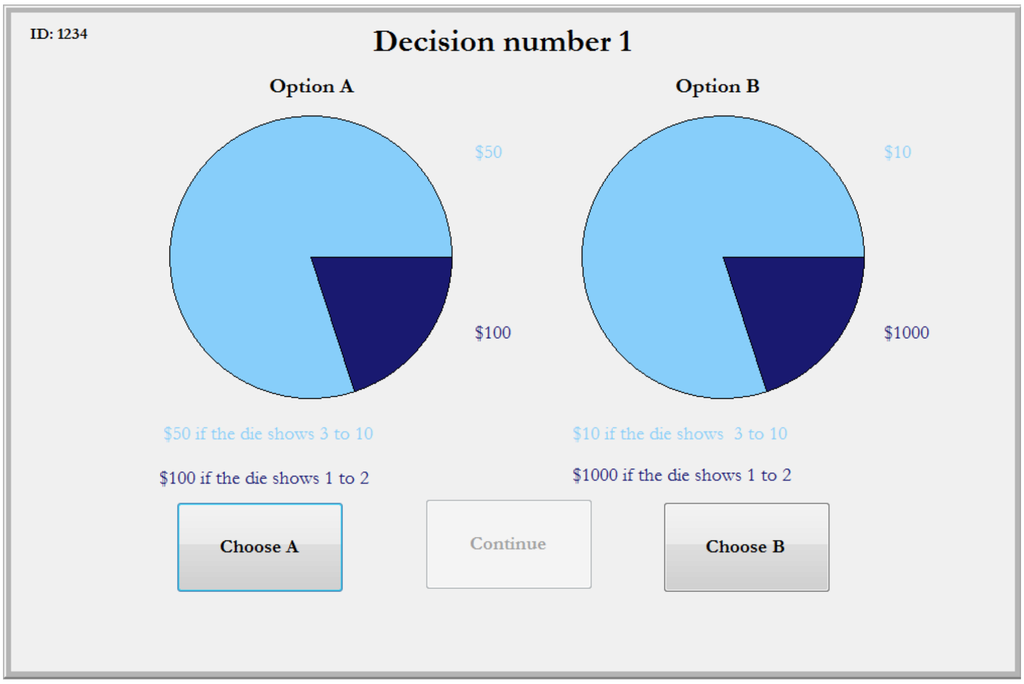

Figure 2: Downtown Network With Bus on 9th Avenue

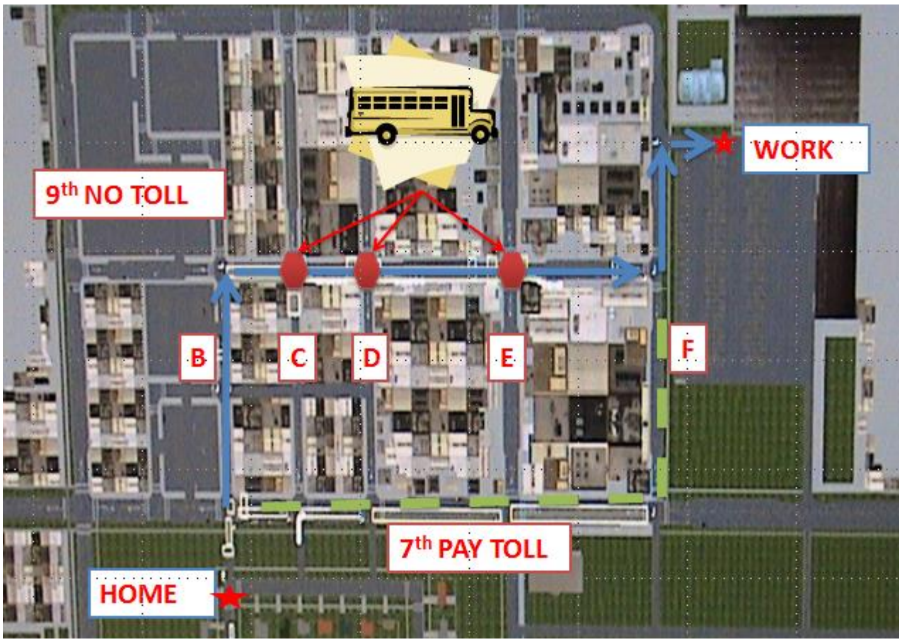

Figure 3: Flow of Tasks in Full Experiment
Tasks Analyzed Here Are Indicated in Grey

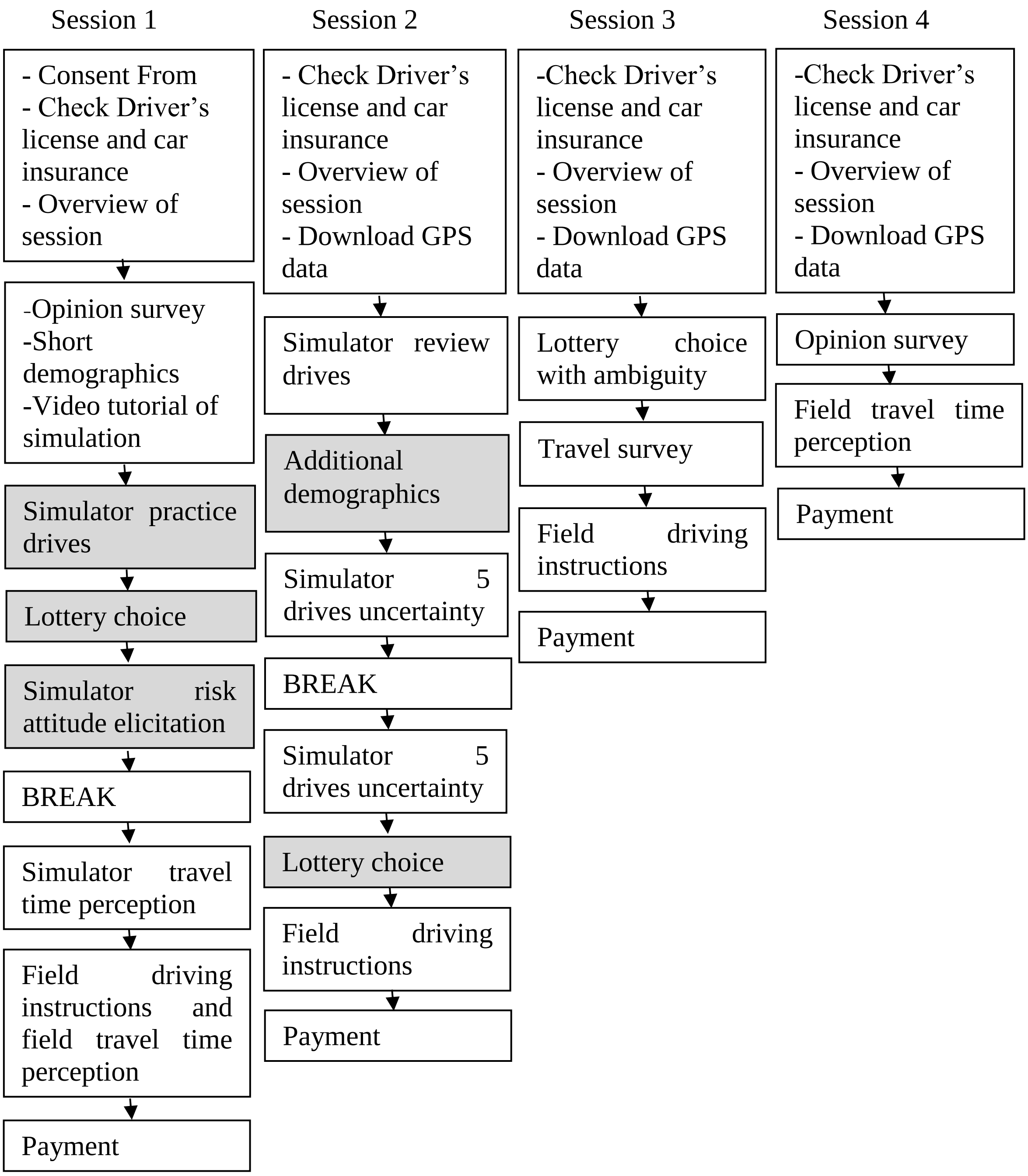

Figure 4: Estimated Distributions of Risk Aversion Coefficients for the EUT CRRA Models

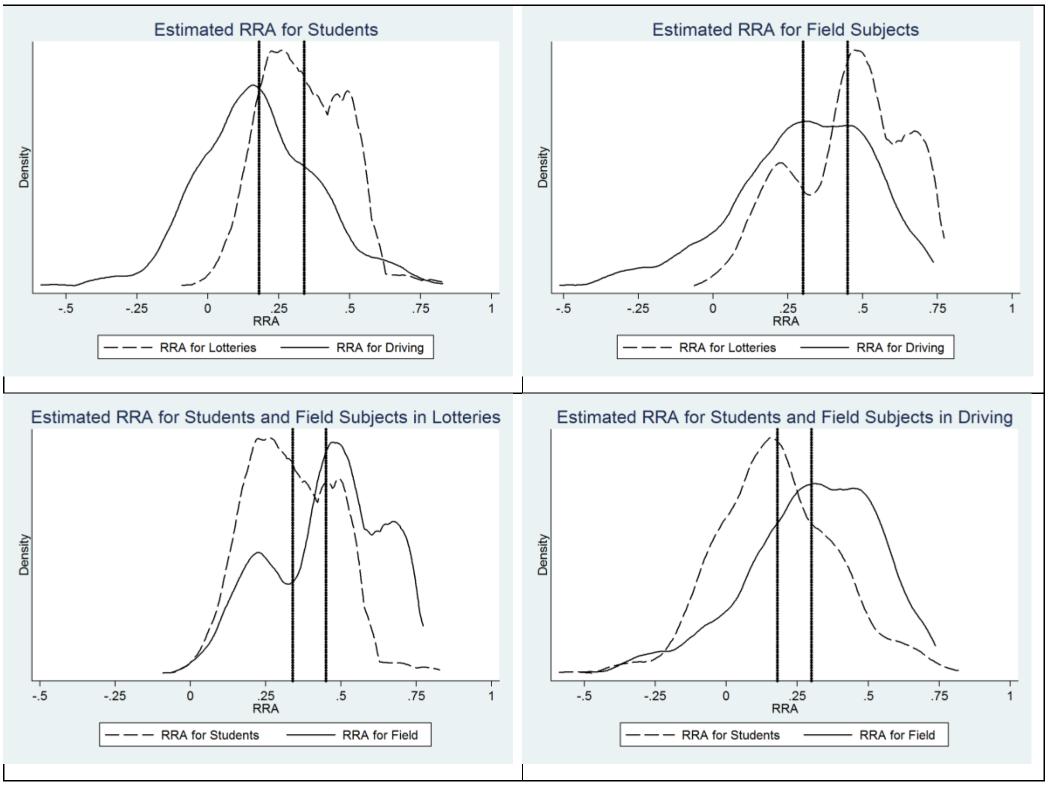

Table 1: Prizes and Probabilities in Lottery Task

Probability range	Safe Lottery Low Prize	Safe Lottery High Prize	Risky Lottery Low Prize	Risky Lottery High Prize
0.1 – 0.9	$2	$3	$0.25	$4
0.1 – 0.9	$2	$3	$0.25	$5
0.1 – 0.9	$2	$3	$0.25	$6
0.1 – 0.9	$4	$6	$0.50	$10

Probabilities were drawn uniformly from the discrete set [0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9]

Table 2: Tolls and Wages in the Simulator Task

	Low Toll Range	Medium Toll Range	High Toll Range
Wage=$4	$0.5-$1.50	$1.60-$2.50	$2.60-$3.50
Wage=$5	$0.5-$1.80	$1.90-$3.20	$3.30-$4.50
Wage=$6	$0.5-$2.80	$2.20-$3.80	$3.90-$5.50

Tolls were drawn uniformly from each set in steps of $0.10.

Table 3: Example of implied pairwise route choice

Probability of bus	Safe Option (7th Ave)	Risky Option (9th Ave) High Payoff	Risky Option (9th Ave) Low Payoff	Expected Value Difference
0.5	$5 - $1.20 = $3.80	$5	0.25	$3.80 - $2.625 =$1.18
0.5	$5 - $1.90 = $3.10	$5	0.25	$3.10 - $2.625 =0.48
0.5	$5 - $3.70 = $1.30	$5	0.25	$1.30 - $2.625 =-$1.325
***Expected Value 9th= 0.5*$5+0.5*0.25=$2.625**				

Table 4: Demographic Characteristics of the Field and Student Participants

	Field	Students
Female	53%	48%
Low Income (<=$50,000)	27%	40%
High Income (>$80,000)	39%	39%
College Degree	78%	na
Age 18_30	26%	95%
Age 56_75	12%	0%
African-American / Black	20%	52%
Frequency of game playing on scale 1- 5	2.04	2.30

Table 5: Exogenous variables

Variable	Description	Min value	Max value
Prob	Probability bad outcome in risky option (congestion probability in driving simulator)	0	1
Prize	High prize in risky option (wage in driving simulator)	$4	$10
Toll	Toll on 7th Ave. in driving simulator	$0.50	$5.50
Task	Counter for task number, 1-4 for lotteries and 1-3 for driving simulator	1	4
Lott2	Indicator for lotteries in Session 2	0	1
CumW	Instrumented cumulative earnings	0	60
WS2	Instrumented cumulative earnings interacted with Session 2 dummy	0	60
Oftenplay	Experience with video and computer games and online worlds: 1=Never, 2=A few times, 3=Occasionally, 4=Often, 5=Every chance I get	1	5
Oldage_play	Experience with video games interacted with being older than 30 years	1	5
Female		0	1
Low Income	Less than US national median household income of $50,000 annually	0	1
High Income	More than $80,000 per year	0	1
College	Completed college degree	0	1
Age 18-30	Excluded age group is 31-55	0	1
Age 56-75		0	1
African-American/Black	All other ethnic groups pooled	0	1

Table 6: Cumulative earnings by task

	Field	**Students**
Lottery task 1	$0	$0
Lottery task 2	$2.99	$3.35
Driving task 1	$6.10	$6.60
Driving task 2	$8.79	$9.66
Driving task 3	$11.72	$12.54
Lottery task 3	$25.41	$26.91
Lottery task 4	$28.29	$30.09

* Since task 6 and 7 are performed at the end of Session2, we include only earnings accumulated during prior tasks in Session 2 for those tasks.

Table 7: Unconditional Proportion of Risky Choices

	Field	**Students**
Lottery task 1	37%	37%
Lottery task 2	43%	47%
Driving task 1	41%	45%
Driving task 2	47%	45%
Driving task 3	47%	51%
Lottery task 3	38%%	42%
Lottery task 4	44%	43%
All tasks	42%	44%
All lottery tasks	40%	42%
All driving tasks	45%	47%

Table 8: Probit Models

	Field lottery	Student lottery	Field simulator	Student simulator	*t*-tests driving (field vs students)	*t*-tests lottery (field vs students)	*t*-tests students (lottery vs simulator)	*t*-tests field (lottery vs simulator)
Constant	.59 (.000)	.74 (.000)	.54 (.001)	.52 (.001)	(.47)	(.003)	(.04)	(.18)
Lott2	-.18 (.10)	-.05 (.66)				(.03)		
Toll			.26 (.000)	.12 (000)	(.000)			
High risky prize	.03 (.006)	.02 (.03)	-.15 (.000)	-.05 (.04)	(.18)	(.34)	(.000)	(.000)
Prob	-.57 (.000)	-.73 (.000)	-.46 (.000)	-.41 (.001)	(.000)	(.000)	(.000)	(.000)
cumW	-.001 (.96)	.02 (.11)	.02 (.006)	.03 (.000)	(.08)	(.003)	(.10)	(.001)
WS2	.003 (.84)	-.02 (.12)				(.002)		
Oftenplay	-.000 (.99)	.01 (.39)	.04 (.15)	.000 (1.0)	(.003)	(.03)	(.04)	(.002)
Oldage_play	-.004 (.91)	-.01 (.92)	.04 (.30)	-.05 (.25)	(.14)	(.06)	(.38)	(.001)
Task	.07 (.08)	.05 (.24)	-.02 (.50)	-.06 (.07)	(.02)	(.30)	(.000)	(.000)

Endogenous variable is the propensity to select the risky option. Numbers in brackets are *p*-values.
Coefficients are marginal effects. Residual errors are clustered on the individual subject.

Table 9: EUT models

	Field lottery	Student lottery	Field driving	Student driving	*t*-tests driving (field vs students)	*t*-tests lottery (field vs students)	*t*-tests students (lottery vs simulator)	*t*-tests field (lottery vs simulator)
r								
Constant	.998 (.000)	.73 (.001)	.75 (.001)	.36 (.14)	(.03)	(.03)	(.02)	(.07)
Lott2	.6 (.04)	-.10 (.72)				(.001)		
cumW	.02 (.72)	-.04 (.25)	-.06 (.007)	-.07 (.004)	(.34)	(.04)	(.07)	(.01)
WS2	-.03 (.67)	.05 (.17))				(.02)		
Oftenplay	-.03 (.63)	-05 (.24)	-.12 (.07)	-.05 (.45)	(.08)	(.33)	(.46)	(.05)
Oldage_play	.03 (.55)	.13 (.41)	.14 (.006)	.08 (.33)	(.15)	(.09)	(.28)	(.01)
Task	-.31 (.004)	-.13 (.25)	.06 (.47)	.28 (.01)	(.01)	(.01)	(.001)	(.000)
μ								
Constant	1.84 (.10)	.90 (.72)	1.005 (.99)	.97 (.92)				
Lott2	1.005 (.99)	.55 (.06)						
CumW	1.04 (.55)	.93 (.12)	1.07 (.03)	.997 (.93)				
WS2	.96 (.55)	1.09 (.11)						
Oftenplay	.90 (.23)	.94 (.36)	.96 (.66)	.97 (.71)				
Oldage_play	1.07 (.33)	.92 (.54)	.90 (.11)	.74 (.000)				
task	.87 (.32)	.19 (.37)	.96 (.77)	.90 (.49)				

Significance for mu is reported as difference from 1. Significance level for r is reported as difference from 0. Numbers in brackets are *p*-values.

Table 10: RDU models

	Field lottery	Student lottery	Field driving	Student driving	*t*-tests driving (field vs students)	*t*-tests lottery (field vs students)	*t*-tests students (lottery vs simulator)	*t*-tests field (lottery vs simulator)
r								
Constant	.64 (.002)	.57 (.03)	.49 (.06)	.46 (.07)	(.44)	(.29)	(.22)	(.19)
Lott2	.13 (.26)	-.18 (.18)				(.001)		
CumW	.01 (.61)	-.03 (.06)	.003 (.70)	-.02 (.13)	(.02)	(.007)	(.18)	(.19)
WS2	-.01 (.60)	.03 (.03)				(.005)		
Oftenplay	-.4 (.37)	-.03 (.14)	-.03 (.21)	.003 (.91)	(.05)	(.42)	(.03)	(.41)
Oldage_play	.03 (.24)	.06 (.24)	-.05 (.76)	-.06 (.04)	(.01)	(.10)	(.001)	(.02)
Task	-.10 (.02)	.01 (.03)	.006 (.87)	.08 (.15)	(.04)	(.002)	(.04)	(.002)
γ								
Constant	1.54 (.001)[a]	1.09 (.001)[b]	1.22 (.000)d	.78 (.000)[c]	(.04)	(.04)	(.04)	(.12)
Lott2	.65 (.13)	.12 (.63)				(.02)		
CumW	-.01 (.93)	-.01 (.84)	-.06 (.03)	na	na	(.49)	na	(.07)
WS2	.000 (1.0)	.01 (.76)				(.39)		
Oftenplay	-.005 (.95)	-.01 (.81)	-.03 (.53)	-.01 (.68)	(.28)	(.46)	(.40)	(.25)
Oldage_play	-.003 (.95)	.07 (.70)	.08 (.12)	-.01 (.68)	(.32)	(.18)	(.36)	(.02)
Task	-.29 (.07)	-.15 (.18)	.07 (.41)	.007 (.90)	(.13)	(.05)	(.005)	(.001)

Significance for γ being different from 1 is a) *p*-value for difference from 1 = 0.24, b) *p*-value for difference from 1 = .74, and c) *p*-value for difference from 1 = 0.28. Numbers in brackets are *p*-values.

REFERENCES

Andersen, Steffen, Vinayak V. Dixit; Glenn W. Harrison, Rami C. Harb, Morten I. Lau, Essam Radwan, Elisabet E. Rutström and Ron Tarr. 2011. "Experiments on Driving under Uncertain Congestion Conditions and the Effects on Traffic Networks from Congestion Initiatives." *Report to U.S. Federal Highway Administration*, Available at: http://dbel.robinson.gsu.edu/files/2013/08/FHWA-experimental-design-and-budget-overview-November-25-2011-v2.pdf

Andersen, Steffen, Glenn W. Harrison, Morten I. Lau, and Elisabet E. Rutström. 2008. "Eliciting Risk and Time Preferences." *Econometrica* 76(3): 583-618.

Andersen, Steffen, Glenn W. Harrison, Morten I. Lau and Elisabet E. Rutström. 2010. "Preference heterogeneity in experiments: Comparing the field and laboratory." *Journal of Economic Behavior and Organization* 73(2): 209-224.

Andersen, Steffen, Glenn W. Harrison, Morten I. Lau and Elisabet E. Rutström. 2010. "Behavioral Econometrics for Psychologists." *Journal of Economic Psychology* 31(4): 553-576.

Anderson, Lisa R., and Jennifer M. Mellor. 2009. "Are risk preferences stable? Comparing an experimental measure with a validated survey-based measure." *Journal of Risk and Uncertainty* 39(2): 137-160.

Barseghyan, Levon, Jeffrey Prince and Joshua C. Teitelbaum. 2011. "Are Risk Preferences Stable across Contexts? Evidence from Insurance Data." *American Economic Review* 101(2): 591-631.

Barsky, Robert B., F. Thomas Juster, Miles S. Kimball and Matthew D. Shapiro. 1997. "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study." *Quarterly Journal of Economics* 112(2): 537-579.

Benz, Matthias and Stephan Meier. 2006. "Do People Behave in Experiments as in the Field? Evidence from Donations." *Working Paper No. 06-8*, Federal Reserve Bank of Boston.

Chaiken, Shelly and Yaacov Trope. 1999. *Dual-Process Theories in Social Psychology*. New York, NY: Guilford Press.

Cooper, David J., John H. Kagel, Wei Lo, and Qing Liang Gu. 1999. "Gaming against Managers in Incentive Systems: Experimental Results with Chinese

Students and Chinese Managers." *American Economic Review* 89 (4): 781-804.

Cox, James C., Vjollca Sadiraj and Ulrich Schmidt. 2011. "Paradoxes and Mechanism for Choice under Risk." *Working Paper 2011-12*, Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University.

Dixit, Vinayak V., Glenn W. Harrison and Elisabet E. Rutström. 2013. "Estimation of Subjective Beliefs of Crash among Drivers." *Working Paper 2013-07*, Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University, forthcoming *Accident Analysis and Prevention*.

Evans, Jonathan and Keith Frankish. 2009. *In Two Minds: Dual Processes and Beyond*. Oxford: Oxford University Press.

Falk, Armin, Stephan Meier and Christian Zehnder. 2013. "Do Lab Experiments Misrepresent Social Preferences? The Case of Self-selected Student Samples." *Journal of the European Economic Association* 11(4): 839-852.

Fiore, Stephen M., Glenn W. Harrison, Charles E. Hughes and Elisabet Rutström. 2009. "Virtual Experiments and Environmental Policy." *Journal of Environmental and Economic Management* 57(1): 65-86.

Fox, Craig R., and Amos Tversky. 1995. "Ambiguity Aversion and Comparative Ignorance." *Quarterly Journal of Economics* 110 (3): 585-603.

Gaudecker, Hans-Martin von, Arthur H.O. van Soest and Erik Wengstrom. 2011. "Heterogeneity in risky choice behavior in a broad population." *American Economic Review* 101(2): 664-694.

George, J. Greg, Glenn W. Harrison, Elisabet E. Rutström and Shabori Sen. 2102. " Behavioural Responses towards Risk Mitigation: An Experiment with Wild Fire Risks." *FUR XV International Conference*, Georgia State University, Atlanta, 30 June 2012- 3 July 2012.

Harrison, Glenn W., Morten I. Lau, and Elisabet E. Rutström. 2007. "Estimating Risk Attitudes in Denmark: A Field Experiment." *Scandinavian Journal of Economics* 109(2): 341-368.

Harrison, Glenn W., Elisabet E. and Rutström. 2008. "Estimating Risk Aversion in the Laboratory." In *Risk Aversion in Experiments*, Research in Experimental Economics 12: 41-196.

Harrison, Glenn W., and J. Todd Swarthout. 2012. “The Independence Axiom and the Bipolar Behaviorist.” *Working Paper 2012-01*, Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University.

Hey, John D., and Chris Orme. 1994. “Investigating Generalizations of Expected Utility Theory Using Experimental Data.” *Econometrica* 62(6): 1291-1326.

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

Huck, Steffen, and Wieland Müller. 2012. “Allais for All: Revisiting the Paradox in a Large Representative Sample.” *Journal of Risk and Uncertainty* 44: 261-293.

Kahneman, Daniel. 2003. “Maps of Bounded Rationality: Psychology for Behavioral Economics.” *American Economic Review* 93(5): 1449-1475.

List, John A. 2007. “Field Experiments: A Bridge Between Lab and Naturally-Occurring Data.” *NBER Working Paper No. 12992*, National Bureau of Economic Research.

Levitt, Steven, and John A. List. 2007. “What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?” *Journal of Economic Perspectives* 21(2): 153-174.

Mukherjee, Kanchan, 2010, “A Dual System Model of Preferences under Risk”, *Psychological Review* 177(1): 243-55.

Offerman, Theo, Joep Sonnemans, Gijs van de Kuilen and Peter P. Wakker. 2009. “A Truth-Serum for Non-Bayesians: Correcting Proper Scoring Rules for Risk Attitudes.” *Review of Economic Studies* 76(4): 1461-1489.

Stanovich, Keith E., and Richard F.West. 2000. “Individual Differences in Reasoning: Implications for the Rationality Debate?” *Behavioral and Brain Sciences* 23: 645-726.

Tversky, Amos, and Kahneman, Daniel. 1992. “Advances in Prospect Theory: Cummulative Representation of Uncertainty.” *Journal of Risk and Uncertainty* 5: 297-323.

Wakker, Peter P., Daniëlle R. M. Timmermans and Irma A. Machielse. 2007. “The Effects of Statistical Information on Risk and Ambiguity Attitudes, and on Rational Insurance Decisions.” *Management Science* 53: 1770-1784.

Wilcox, Nathaniel T. 2011. “Stochastically More Risk Averse: A Contextual Theory of Stochastic Discrete Choice under Risk.” *Journal of Econometrics* 162(1): 89-104.