MEASURING RISK AVERSION TO GUIDE TRANSPORTATION POLICY: CONTEXTS, INCENTIVES, AND RESPONDENTS

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

April 2015

ABSTRACT Road pricing may provide a solution to increasing traffic congestion in metropolitan areas. Route, departure time and travel mode choices depend on risk attitudes as commuters perceive the options as having uncertain effects on travel times. We propose that Experimental Economics methods can deliver data that uses real consequences and where the context can be varied by the researcher. The approach relies on the controlled manipulation of contexts, similar to what is done in the Stated Choice approach, but builds in actual consequences, similar to the Revealed Preference approach. This paper investigates some of the trade-offs between the cost of conducting Experimental Economics studies and the behavioral responses elicited. In particular, we compare responses to traditional lottery choice tasks to the route choice tasks in simulated driving environments. We also compare students (a low cost convenient participant pool) to field participants recruited from the driving population. While we see initial differences across our treatment groups, we find that their risk taking behavior converge with minimal repetition.

KEYWORDS Congestion pricing, risk attitudes, contextual tasks, experiments, driving simulations

---

[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, erutstrom@gsu.edu (Rutström). All online materials available at http://dbel.robinson.gsu.edu/research-projects-fhwa/. 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

In the area of transportation policy an important issue is how congestion pricing may provide a solution to the increasing degree of traffic congestion in most metropolitan areas. To address this question it is important to understand behavioral responses to congestion pricing when travel times are unreliable. Such behavioral responses depend on the distribution of risk attitudes in the driving population. The majority of studies on the willingness to pay for travel time savings rely on assuming risk neutrality, although since the mid-1990s we have seen an emergence of academic research with a focus on the role of risk aversion in transportation choices (see Senna 1994; Bates, Polak, Jones, and Cook 2001; Brownstone and Small 2005; Small, Winston and Yan 2005; Bhat and Sardesai 2006; Hensher, Greene, and Li 2011; and Devarasetty, Burris and Shaw 2012). Despite the two decades that have passed, Rasouli and Timmermans (2014, p. 79) express the concern that "[t]he overwhelming majority of models in travel behavior research assume implicitly or explicitly that individuals choose between alternatives under conditions of certainty." Li, Hensher and Rose (2010) also found that in project appraisals the value of travel time reliability have been ignored, but De Jong and Bliemer (2015) report that several countries have begun to adopt some measures of travel time reliability, although using relatively simple methods. Ignoring or using only poorly measured values of travel time reliability could lead to significantly biased assessments. Fosgerau and Karlström (2010) report that about 15% of the travel time cost would be unaccounted for if not including the value of travel time unreliability. Therefore, more attention to risk attitudes is warranted.

The evidence of risk aversion in travel choices comes predominantly from studies using the Stated Choice (SC) approach. In this approach researchers present a large number of hypothetical driving scenarios to respondents, and ask them to make route or departure time choices in each scenario. The scenarios usually differ in travel time and road pricing. The SC approach is low cost since it involves only surveys of intentions, and can identify a large set of decision model parameters. Although the SC methodology can be informative, its main limitation derives from the absence of real consequences to the choices expressed in the surveys. This lack of consequences can lead to not just a great deal of noise in the responses, but also, more seriously, to response biases (Holt and Laury, 2002; Cummings, Harrison and Rutström, 1996; Li, Hensher and Rose, 2010). A common way in recent SC studies to decrease such biases is to relate the hypothetical scenarios presented to scenarios that participants reveal as being familiar, such as usual commuting routes and their attributes (Hensher, 2010), or to use cheap talk to nudge respondents to be aware of response biases or to use certainty scales (Fifer, Rose and Greaves (2014) .

The Revealed Preference (RP) approach provides an alternative to the SC approach with real consequences. The RP approach involves directly observing the choices of drivers in the field, with naturally occurring consequences such as variations in travel time and road pricing. The limitation of the RP approach is that it cannot implement a large set of contexts since it relies on existing field contexts (Louviere, Hensher and Swaite, 2000). Further, when the interest is in characterizing risk attitudes RP data may be confounded by unobserved variations in the perceptions of travel times and other consequences.

We propose that the Experimental Economics (EE) approach can deliver data that uses real consequences and where the context can be varied by the researcher. The EE approach relies on the controlled manipulation of contexts, similar to what is done in the SC approach, but builds in actual consequences (Harrison and Rutström 2008). The majority of this literature is based on non-contextual choice situations, such as choices over various lotteries that differ in probabilities and prizes, but examples with field contexts can also be found (List and Lucking-Reiley, 2000; Fiore, Harrison, Hughes, and Rutström, 2009; Dixit, Harrison and Rutström, 2014). The drawback of the EE approach is that it can become expensive if a large set of responses is needed, since the cost of the consequences built into the tasks will add up. This paper investigates some of the trade-offs between the cost of conducting EE studies and the behavioral responses elicited.

The value to transportation planning of estimating risk attitudes is to generate better predictions of behavioral responses. Jackson and Jucker (1982) appear to be the first empirical application of risk attitudes in transport, using a mean-variance utility approach. Senna (1994), further investigated this by adding a separable travel cost variable to the model, and found both risk averse and risk preferring responses using SC data. Li, Hensher and Rose (2010) compared estimates of value of travel time and value of travel time reliability across several empirical papers using SC or RP data. Generally, their review revealed that participants are willing to pay both for reduced mean travel time, and for reduced travel time variation, or for the likelihood of arriving late or early (as in scheduling models), consistent with risk aversion. They found that the ratio of the value of travel time reliability to the value of travel time savings vary greatly. Small, Noland, Chu, and Lewis (1999) report values for mean travel time at \$3.90 and values for travel time variability at \$12.60 using a mean-variance utility approach, and the SC data implying that the latter is over three times that of the former. Asensio and Matas (2008) also report values with similar patterns. However, Small, Winston and Yan (2005) find the opposite relationship when using both SC and RP data. They report values for mean travel time at \$12 and \$21.50 for the SC and the RP data, respectively, and values for travel time variability at \$5.40 and \$19.70 for the SC and the RP data, respectively. Hensher, Greene, and Li (2011), further, find some risk loving behavior based on SC data from Australian participants. They conclude that 66% of their participants are risk loving and 34% are risk averse, thus providing evidence of heterogeneity in risk attitudes. Carrion and Levinson (2012) while reviewing evidence literature also find great variation in the ratio of the value of reliability to value of time. They report that the ratio range from 0.10 - 2.51, with Ghosh (2001) and Yan (2002) reporting RP estimates to be higher than the

SP estimates. Li, Hensher and Rose (2010) attribute much of the variation in inferences about risk attitudes to the variations in how information is presented to participants.

Most of this evidence relies on Expected Utility Theory (EUT). In this paper we estimate both EUT models and Rank Dependent Utility (RDU) models, but do not include loss aversion as in Prospect Theory (PT). According to Li and Hensher (2011), very little evidence has been collected in the transport literature that is suitable to testing Prospect Theory. According to the PT literature and its applications on SC data (Chateauneuf and Cohen, 1994; Attema, Brouwer and L'Haridon, 2013) loss aversion and risk loving preferences in the loss domain are important behavioral considerations. However, Li and Hensher (2011) give an overview of the requirements necessary to implement PT, and find very few studies in transport that fulfill these. The most important difficulty in applying PT is the specification of a credible reference point. Avineri (2006) demonstrate how PT valuations are sensitive to the choice of reference point. Individual drivers' reference points are bound to vary with their experiences and expectations, and are difficult to measure jointly with their preferences. Stott (2006) evaluated the various functional forms for the utility function and probability weighting. Note that loss aversion is not required for changes in utilities over losses to loom larger than changes in utilities over gains: that is accomplished already from the concavity of the utility function.

One possible explanation to why SC findings vary with the presentation format is the lack of clear incentives that may affect the individual's incentives to choose according to his/her actual preferences as well as attention and comprehension. It is reasonable to expect that participants pay more attention to the task and think harder about choices and their consequences when incentives are directly linked to the consequences. Holt and Laury (2002) test this hypothesis directly and find that high hypothetical incentives result in similar inferences about risk attitudes as low real incentives, but that significantly more risk aversion is found when the real incentives are scaled up. Cummings, Harrison and Rutström (1995) demonstrate that lack of incentives in valuation studies lead to hypothetical bias, even when valuations are over relatively familiar goods. Evidence of risk aversion is strong across SC, RP and EE studies, but the absence of risk loving behavior is particularly strong in the EE literature, where all studies use monetary incentives. Harrison and Rutström (2008) review several EE studies that vary in many aspects of their approach and conclude that: "… there is systematic evidence that subjects in laboratory experiments behave as if they are risk averse. Some subjects tend towards a mode of risk neutrality (RN), but very few exhibit risk-loving behaviors. The degree of risk aversion is modest, but does exhibit heterogeneity that is correlated with observable individual characteristics." (p. 43)

When monetary incentives are used, as in EE studies, while there is little evidence of risk loving behavior, there is a great degree of heterogeneity in risk attitudes across individuals (e.g., Andersen, Harrison, Lau, and Rutström (2008) (2010), Gaudecker, Soest, and Wengstrom (2011)). It is reasonable to expect that the distribution of risk attitudes may vary across transportation regions so that it may be necessary to carry out estimations of risk attitudes for each individual congestion pricing project. Because this could lead to concerns about costs using the EE approach, in this paper we compare a lower cost EE approach to a higher cost EE approach. First, we compare the common non-contextual binary lottery choice that requires no expensive equipment to carry out, to a task that adds complexity and context by placing a real-time binary route choice in a computer simulated driving task. Second, we compare the responses by a field driving population to those of a convenient, low-cost population of university students.

The purpose of comparing the relatively inexpensive lottery task to the more costly simulated driving task is to test if the elicited risk attitudes are invariant to how the choice is presented to the respondent. There is evidence that variations in presentation formats can lead to variations in elicited risk attitudes. See for example Li, Hensher and Rose (2010), Andersen, Harrison, Lau, and Rutström (2008), Reynaud and Couture (2012) Deck, Lee, Reyes, and Rosen (2013, 2014). Such behavior is consistent with dual process theories in psychology (Chaiken and Tope [1999], Stanovich and West (2000), Kahneman (2003), Evans and Frankish (2009)). Unfamiliar tasks, such as the lottery tasks when presented to field participants as we do here, may involve slower cognitive modes that rely on explicit deliberation, and the simulated driving task that has more familiar elements may involve faster cognitive modes involving emotions and heuristics, so called 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.

One way to approximate the contextual relevance (and familiarity) of RP studies, but at a much lower cost, is to use driving simulators, or computer simulated driving environments. As is the case in the field, the participant is actually engaged in a real-time driving task while making the route choice, compared to in a lottery choice task where participants are not engage in any other simultaneous tasks. The feedback regarding the outcome of the choice, whether lucky or not, is through visually experiencing the traffic congestion, and reacting to it by having to slow down. Thus, even when a simulated driving environment is not modeled on the participant's natural driving environment, the process of driving and making route choices is likely more familiar than the process of choosing between options in experimental lottery tasks.

Consistent with recent findings in both SC and RP studies, our empirical analysis demonstrates that it is important to worry about risk attitudes for transportation policy analysis. We find strong evidence of significant risk aversion, with virtually no risk loving behavior after minimal experience with the tasks. Our findings are optimistic about the use of low cost methods. We find that the lower cost methods (binary lotteries and student participants) generate data that is consistent with the higher cost methods (simulated driving tasks and participants from the general driving population) if participants are given an opportunity to learn through experience. The risk preferences inferred from the low cost methods match those from the high cost method, but only after participants have had some experience with the consequences of their choices. The demonstrated effect of experience imply that incentives alone may not be sufficient to generate robust measures of risk attitudes, but that attention to the presentation format and the ability to learn through experience is also necessary.

2. THEORY

All the tasks we present to our study participants consist of binary choices over risky outcomes with known, objective probabilities. Each task involves a safer and a riskier option, where each option has two possible outcomes, one which is better than the other. Following much of the literature in EE and a growing literature in transportation our initial theoretical approach is based on Expected Utility (EU) maximization and a Constant Relative Risk Aversion (CRRA) utility function. There is a large literature in economics, finance and psychology applying and evaluating various choice theories, utility functions, and probability weighting functions. The CRRA utility function, also known as the power function, and the CARA function, also known as the exponential function, are both popular in this literature. Stott (2006) examines both of these in the context of PT and Wakker (2010: section 3.5.1) compare their properties. Examples of applications to EUT in the EE literature include Holt and Laury (2002), several studies reviewed in Harrison and Rutström (2008), and Hey and Orme (1994), who provide empirical arguments in favor of EUT over several non-EUT alternatives. In transportation, the EUT framework (in either a mean-variance formulation or a scheduling formulation) was adopted by Small, Winston and Yan (2005), Bates, Polak, Jones and Cook

(2001), Hollander (2006), Asensio and Matas (2008), Senna (1993), and Li, Hensher and Rose (2010). Hensher, Greene and Li (2011) and Dixit, Harrison and Rutström (2013) adopt a similar non-linear approach to what is used here.

The EUT approach can easily be extended to Rank Dependent Utility (RDU) or Prospect Theory (PT) with a variety of utility function specifications. In otherwise similar modeling frameworks, Hensher, Greene and Li (2011) use RDU and Li and Hensher (2011) apply PT to SC data. Stott (2006) compare a large number of utility functions, probability weighting functions, and stochastic choice models on SC data within the PT framework. Hey and Orme (1994) and Wilcox (2015a, b) use non-parametric utility functions and compare several decision model specifications across EUT, RDU and PT, fitting models to a subset of the observations and predicting on the remainder based on EE data. Harrison and Rutström (2009) offers an alternative to classifying each subjects by applying a statistical mixture model to estimate relative support for EUT and PT. Wilcox (2011) explores the role of heteroskedastic errors in several latent decision models with important implications for controlling for the outcome context. A somewhat different approach is the polynomial-additive utility representation which is evaluated under various SC outcome contexts in Davis-Stober and Brown (2013). Here we rely on EUT but provide estimates of RDU as well.

The popular mean-variance utility model can be shown to correspond to EU maximization with a Constant Absolute Risk Aversion (CARA) utility function and with travel times distributed according to a normal distribution (Sargent (1987, pp. 154-155)). In the transport literature the motivation for the use of the mean-variance utility function has been independent of its relationship to CARA utility functions, and the standard deviation of travel time is frequently used instead of the variance (Jackson and Jucker (1982). We therefore see the mean-variance utility model as a special case of the more general approach we use here. The EU of option i, where i=*{safe, risky}*, is given by:

$$EU_i = p\,u\big(x_{i,A}\big) + (1-p)\,u(x_{i,B}) \tag{1}$$

where p is the probability of getting outcome x_{iA}. Applying the CRRA utility function for outcome j, where $j=\{A, B\}$, is:

$$u(x_{ij}) = \frac{x_{ij}{}^{1-r}}{1-r} \tag{2}$$

where r is the coefficient of relative risk aversion, defined by Arrow (1965) and Pratt (1964) as

$$r = -\frac{x_{ij}u''}{u'} \tag{3}$$

The utility function is denoted by u, and we define both the first-order (u') and second-order (u'') derivatives as functions of x_{ij}, the argument of the utility function. 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.

To give the model some context, consider the case of a discrete route choice in the field, such as in an RP study. We will simplify here by assuming that the continuous travel time distribution in the field setting can be represented with the binary probability distribution used in equation (1), and only focus on the argument in the utility function, x_{ij}. A simple, linearly separable, specification of this argument would assign each trip a trip value, *v*, and a value of time, *w*, neither of which varies across the route options, *i*. A road price, T_i, is subtracted from the net value and does vary across routes. The CRRA utility function would then be defined over the argument that represents the net value of taking route *i* and experiencing travel time outcome t_{ij}, where again *j*={*A,B*}:

$$u(x_{ij}) = \frac{(v - wt_{ij} - T_i)^{1-r}}{1-r} \tag{4}$$

The EU for route *i* can then be expressed as

$$EU_i = p\frac{(v - wt_{i,A} - T_i)^{1-r}}{1-r} + (1-p)\frac{(v - wt_{i,B} - T_i)^{1-r}}{1-r} \tag{5}$$

A driver who is approaching a location where a choice over routes must be made would evaluate the EU for each route *i*, and choose the route with the maximum EU.

This approach can be generalized in several ways. For example, to implement a mean-variance utility approach, which is more familiar to some readers, a CARA utility function would replace the CRRA function in equation 4: $u(x_{ij}) = -e^{-\lambda(v- wt_{ij}-T_i)}$, and after assuming that t_{ij} follows a continuous standard normal distribution we would derive the mean-variance utility function as $EU_i = -e^{-\lambda\left(w\mu_i - \frac{\lambda w^2 \sigma^2{}_i}{2} + (v-T_i)\right)}$, where μ_i is the mean of the standard normal travel time distribution and $\sigma^2{}_i$ is the variance. Apart from the possibility of using more flexible utility functions it also generalizes to Rank Dependent Utility Theory (Quiggin 1982) and Prospect Theory (Kahneman and Tversky 1979, and Tversky and Kahnemann 1992). Since the distinction between RDU and PT are only relevant when the choice sets include both gain and loss domains, they are equivalent for our purposes here. Lapparent (2010) used an RDU framework to study air travel behavior, and Hensher, Greene and Li (2011) used a variation of RDU for route choices.

The main difference between EU and RDU (as well as PT) is that preferences over risk in the latter can be expressed not only through the utility function but also over a probability weighting function. Using a simple power weighting function we then have the following transformation:

$$W_A = p^\gamma \text{ and } W_B = 1 - p^\gamma \tag{6}$$

where W_j is the decision weight given to outcome *j,* and *j={A,B}.* There are many functional forms for W_j in the literature, but Wilcox (2015b) shows that a concave function (defined over the probability of the best outcome), corresponding to $\gamma < 1$ when the probability *p* is the likelihood of the best possible outcome, is the most common one. Using our binary route choice example from equation (5) we will then have an RDU valuation as:

$$RDU_i = p^\gamma \left(\frac{(v- wt_{iA}-T_i)^{1-r}}{1-r}\right) + (1-p^\gamma)\left(\frac{(v- wt_{iB}-T_i)^{1-r}}{1-r}\right) \tag{7}$$

The simple power weighting function can be given a nice behavioral interpretation. If $\gamma<1$, for example, then $W_A>p$ and the function is everywhere concave. Thus, the agent puts more weight on the likelihood of outcome *A* happening than what is implied by *p*. When *A* is an outcome that is preferred to *B*, then this implies optimistic preferences. On the other hand, if $\gamma>1$, then

$W_A < p$, the function is everywhere convex, the agent puts less weight on the likelihood of outcome *A* happening than what is implied by *p*, and the implied preferences are pessimistic. Pessimistic preferences mean that outcome *A*, the assumed preferred outcome, carries less weight in the RDU valuation than it does with neutral or optimistic preferences. Thus, pessimistic preferences expressed through probability weighting have a similar effect on the valuation of the route options as a concave utility function and can thus be considered another aspect of risk aversion. It is possible for agents to be risk averse in both utility and probability weighting, risk loving in both, or risk averse in one and risk loving in the other (Chateauneuf and Cohen 1994).

3. EXPERIMENTAL DESIGN

Each participant is presented with four lottery tasks and three route choice tasks in the simulated traffic environment. Following our theory exposition, all choice options are binary, and all options have two possible outcomes.[1]

3.1 Computer Simulated Route Choice Task

We include the computer simulated driving task as a way to generate a driving context that is closer to that found in RP studies. The main aspect from an RP design that we incorporate in the simulated task is the fact that the participant is driving while making the route choice, compared to in a lottery choice task where participants are not engage in any other simultaneous tasks. The feedback regarding the outcome of the choice, whether lucky or not, is through experiencing the traffic congestion, and reacting to it by having to slow down.

[1] Task instructions are available online as Appendix A in the supporting document to DBEL-WP1301 at http://dbel.robinson.gsu.edu/research-projects-fhwa/. The tasks analyzed here are part of a larger study described in Andersen, Dixit, Harrison, Harb, Lau, Radwan, Rutström and Tarr [2011].

There are, however, limits on how well the simulated environment can represent field attributes. While the simulation software can deliver variations in travel time it is questionable whether the value of the travel time in a simulated environment is informative of the value of travel time in field settings. The value of time is given by the opportunity cost of that time, in addition to any late or early arrival penalties. In a special research session, however, the participant has set aside a total time for participation. Once in the session, during a simulated driving task the opportunity cost of the travel time is measured only in activities that could have been undertaken as part of the research session, not those that could have been undertaken <u>instead</u> of being in the research session.

The solution to the problem of representing appropriate values in a research session is to induce them: to directly control for their influence on behavior by varying them in known ways (Smith, 1976). The reward medium used to induce value of time must fulfill the same properties as the value of time itself. Thus, since utility is a monotone decreasing function of the value of travel time it must also be true for the reward medium. Second, the reward medium must dominate any other motivations that the participant may perceive to be present. One reward medium that has these properties is money. Thus, instead of relying on the inappropriate subjective value of time in the simulated task, we impose a monetary cost on the delay time.

Another factor in the simulated task that can confound our inferences on risk attitudes is variations in travel time due to participants managing the simulated driving task differently. Participants that have a harder time driving in the simulated environment may choose to drive slower even in free flow traffic, thus decreasing the travel time difference between non-congested and congested scenarios. If not controlled for, this can make them appear less risk averse than they are since the travel time delay due to congestion on the risky route is less than it is for participants who have an easier time in the simulated task. To deal with this problem the only travel time consequence that we include in the utility argument is an exogenous travel time difference across routes and across congestion conditions. The term wt_{ij} in the utility function in equation (4) therefore does not vary with such idiosyncratic travel time variations, but is fixed for each participant. It does, however, vary with route and congestion condition.

The simulated driving task presents each driver with a choice between a safer and a riskier route option, where both have monetary consequences (representing the lost value of time). In the instructions to the simulated driving task, participants are 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. The options consist of two routes that each driver can take when driving between home and work: *7th Ave.* is a route that does not get congested but where a toll is charged, and *9th Ave.* is a route that can get congested due to the random presence of a school bus but where a toll is not charged. Thus, *7th Ave.* is the safe option and *9th Ave.* is the risky option. The difference from the lottery task is that there is no congestion risk at all on *7th Ave.* Both of these routes go through the simulation's downtown area.[2]

Participants are initially introduced to the simulated task via a map of the downtown area, shown in Figure 2. They only see the map perspective during an introduction video.[3] During the choice tasks the participant views everything from a perspective similar to that of sitting in a car, which gives them a sense of familiarity. Since the simulated area is not a replication of actual routes in Orlando or Atlanta, which are the cities that our participants are from, the only familiarity with the routes themselves comes from three practice drives

[2] The driving simulation software 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.

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

and the introduction video.[4] The driver's car is initially parked on *B Street* just south of the intersection with *6th Ave.* (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 Ave.* (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 we have added simulated vehicles that drive at the speed limit. These added vehicles also increase the realism of the simulated environment. The choice task is binary: they can only take *7th Ave.* or *9th Ave.* 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 Ave.* 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 Ave.* or to continue straight to take *9th Ave.*. The congestion risk on *9th Ave*. 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 Ave.* 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.

Each participant makes three practice drives and three paid drives in the simulated environment. The first practice drive requires them to take *7th Ave.*, and the other two to take the *9th Ave.*, the first time without a school bus and the

---

[4] This is an important difference between our simulated drive task and an RP study, or an SC study that uses a revealed preference option as a reference point. However, the purpose here is to investigate relatively inexpensive ways to elicit risk attitudes, while imposing real consequences in controlled scenarios. If responses are invariant to the variations we introduce here, given the incentives we impose and the controls we use, then there is reason to be optimistic about them measuring the risk attitudes that guide field behavior as well.

second time with a school bus. Therefore they have experienced each of the three possible drives.

We model the simulated route choices as in *Eqs. (4)-(5)* for EU and (*6)-(7)* for RDU. *9th Ave.* is the risky option because there is a probability that drivers will encounter a school bus, resulting in a loss in value of time, which is induced using money as explained earlier. This probability varies across participants and can be 0.3, 0.5 or 0.7. Thus, the probability of getting the preferred (non-congested) outcome is one minus this probability. 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 participant 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 participants.

Probabilities and monetary outcomes are assigned randomly to participants but are fully known before any choices are made. The trip value, v, can be \$4, \$5 or \$6 and varies across participants, but stays the same for the three drives. If there is congestion due to a schoolbus the induced value of time lost (due to congestion delay) on *9th Ave.* (wt_{ij}) is \$3.75, \$4.75, and \$5.75, respectively for each of the trip values. There is no value of time lost due to congestion delays on *7th Ave.* since it is never congested, however, there is a toll to be paid, T_7. The toll is known to the driver before each drive, but varies across the three drives and across drivers. The toll is randomly 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 trip values so that we avoid the possibility that the toll exceeds the trip value (since we cannot have a participant pay out of pocket). We vary the order in which drivers encounter the low, medium and high range. 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 trip value is \$5, and the value of time lost when ending up behind a bus is \$4.75 leading to a net payoff of \$0.25. This implies that the participant is facing a choice of *9th Ave.* that pays \$5 with a 50% chance and \$0.25 with a 50% chance or a choice of *7th Ave.* that pays \$5 minus the toll. 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 Ave.* exceeds that of taking *9th Ave.* for the two lower illustrative tolls, so a risk neutral driver would be expected to choose *7th Ave.* for all but the highest of these three toll values (\$3.70). However, a sufficiently risk averse driver will choose to take *7th Ave.* even when the expected value of taking *9th Ave.* exceeds the expected value of taking the other route. In this way we can draw inferences about risk attitudes.

3.2 Binary Lottery Task

The less expensive task is a classic context-free EE lottery choice design (Hey and Orme, 1994; Holt and Laury, 2002). 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 same probability of getting the high prize (the preferred outcome) is applied to both lotteries, but it varies across tasks. The left lottery is always the relatively safe one (with a smaller difference between the prizes), and the right lottery is always the relatively risky one (with a larger difference between the prizes). We use ten-sided dice to implement the random process and assign lower number outcomes on the die to the low lottery prize. By varying how many sides of the die yield the lower prize, we vary the probability in a way that is transparent to the participant. 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 hypothetical stakes that are very large to make it clear that it is just practice. No payments are made for the practice task. The program randomly selects a combination of prizes and probabilities from Table 1 for each paid task, 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 a ten-sided die to play out the selected lottery and the earnings are recorded, along with the cumulative earnings.

We model the valuation of the lottery choices as shown in *Eqs. (1) – (3)*. The argument x_{ij} in the utility function is simply the lottery prize, and the probability, p, is applied to the higher of the two prizes.

3.3 Participant Pool Treatments

Apart from investigating behavioral differences due to the presentation format of the choice we also test if participants recruited from the general driving population differ in important ways from participants recruited from a more convenient and less expensive population: university students. The latter group is commonly recruited for a wide range of economics experiments, but when the purpose of a study is to guide policy it is important to ask how representative the risk attitudes are of the affected population. Andersen, Harrison, Lau and Rutström [2010] report that students predict field participant behavior in lottery tasks at least within the demographic ranges represented by the students. However, behavior may differ across these participant pools depending on their familiarity with the task. 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 participants in stylized task settings. 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. 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 participant behavior may differ as a function of familiarity with the environment. In our case, we expect a great deal of heterogeneity in the responses but it may still be possible for a student population to have some predictive power for the general population.

We recruit participants from the general population in Atlanta, Georgia and Orlando, Florida, and students from the Georgia State University. All sessions were carried out from June 2011 until November 2012.

3.4 Recruiting and Experimental Procedures

The tasks analyzed here are part of a larger experiment described in Andersen, Dixit, Harrison, Harb, Lau, Radwan, Rutström and Tarr [2011].[5] In the larger experiment field participants come to four sessions separated by approximately two weeks each. Two of the lottery tasks and all three driving tasks are conducted during Session 1, and the final two lottery tasks in Session 2. Within each session tasks in the simulated environment are alternated with non-simulated tasks to minimize the chances of simulation induced nausea. Participants arrive individually in a staggered manner and go through the series of tasks sequentially.

Instructions are read out individually, and an interviewer is assigned to assist the participant in each task. In the first session participants are first shown a video introducing them to the simulated environment and the routes they may choose from. After that they are able to make three practice drives, corresponding to each of the three driving conditions they may encounter in the actual paid tasks. The first task that is not performed in the simulated traffic environment, immediately after these practice drives, is the binary lottery task. They are given one practice lottery, followed by two paid lottery choices. The next immediate task consists of three paid simulated drives. Participants then complete several additional tasks in Session 1 that are not of relevance here.

[5] Instructions can be found online in supporting documents to DBEL-WP1301 at: http://dbel.robinson.gsu.edu/research-projects-fhwa/

Returning to Session 2 they complete other paid tasks before finishing that session with an additional lottery task, also consisting of two paid lottery choices. 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. The earnings tracking provides the participants with important feedback that influences their experience and familiarity with the tasks. We control for the influence from their cumulative earnings 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. This payment protocol has both theoretical (Cox, Sadiraj and Schmidt 2011) and empirical problems (Harrison and Swarthout 2012), particularly if estimating RDU models, in addition to generating a lot of behavioral noise.

Field participants were recruited by invitation letters. Recipients were randomly selected from the United States Postal Service (USPS) mailing lists, with oversampling from mail carrier routes that have median income levels below the state-wide median income 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. 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 of this study. 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 computer simulated environments 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. About 140,000 letters were sent out which resulted in 633 participants who showed up to the first meeting. This unusually low response rate is likely due both to the inclusion of many mail carrier routes with especially low income levels and to our requirement that participants must communicate with us via email. In the U.S., recruiting low income participants to research studies using regular mail invitations is often more difficult and lower income participants are also less likely to have easy access to the internet. Out of these 283 completed all the lottery and simulated driving tasks analyzed here. The remaining participants were either given a different set of tasks, or (a small portion) dropped out between the first and second session. Our retention rate between Session 1 and Session 2 is 91 percent. In addition we recruited students on the campus of Georgia State University during the summer and fall 2012. From this effort we obtained complete participation by 205 students.

We collected choice data from 5 cohorts of field participants: 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. 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 participants and the students. Since all the students are in college, none of them has yet a college degree. 22% of our field participants have not completed a college degree. The gender distribution is similar across field and student participants, and so is the income distribution, but otherwise there are some differences, as would be expected. The income variable we use 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 close to that 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.

4. EMPIRICAL ANALYSIS

We want to test if students are different from field participants in terms of their risk attitudes as elicited from the lottery task and the driving simulation task. We also want to test if behavior in the simulated driving environment is different from behavior in the lottery tasks as elicited from both field participants and students.

We will first describe the estimation approach and describe the raw data of proportions of choices in which the risky option was chosen. We then estimate simple Probit models to condition the choices on treatment variables and demographic characteristics. Finally, we estimate structural non-linear decision models that will allow us to draw inferences about risk attitudes. In particular, we estimate both an EU model, as shown in *Eqs. 1-5*, and an RDU model as shown in *Eqs. 6-7*. Behavioral differences are captured by the risk attitudes, i.e. the curvature of the utility functions and the probability weighting functions. We use Maximum Likelihood techniques and estimate the full nonlinear models, rather than reduced forms. A similar approach was used in Hensher, Greene, and Li (2011), who also use CRRA utility functions and also estimate an approach similar to RDU. Others using similar non-linear specifications include Gao, Frejinger, and Ben-Akiva (2010), and Dixit, Harrison and Rutström (2013). Stott (2006) compares a large number of utility and probability weighting specifications using a similar approach, and Wilcox (2011, 2015a and b) provides discussions of various error structures.

4.1 Estimation approach

Our econometric strategy is based on Full Information Maximum Likelihood (ML) estimations of the latent decision models defined in equations *1-7*. All observations are binary choices of routes or lotteries. We use a standard normal probability distribution in the ML model defined over the choice across the two options that differ in riskiness:[6]

$$\phi(\varphi) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\varphi^2} \tag{8}$$

where φ is

$$\varphi = \Delta EU = (EU_{risky} - EU_{safe})/\mu \tag{9}$$

the difference in EU (from *eqs. 1-5*) across options *risky* and *safe*, scaled by a parameter μ. This parameter has a behavioral interpretation: it is the participant's sensitivity to the valuation differences, and is also referred to as Fechner error. When μ=0 the agent is infinitely sensitive and the model essentially becomes deterministic. Any value μ>0 implies that the choice is stochastic, and the larger is μ the more random the modeled choice is.

Fechner (1860) studied the physiological sensitivity in humans, and the application of this sensitivity parameter in decision theories has been popularized by Becker, DeGroot and Marschak (1963) and Hey and Orme (1994). Hensher, Greene and Li (2011) refer to this as a random scaling parameter. We further transform (9) following the contextual utility approach by Wilcox (2011 and 2015a), who demonstrated that when choice probabilities are given by value differences such as in Eq. (9), it requires the use of stochastic choice models that are based on heteroscedastic errors and offers Contextual Utility as one such model.

Recall that the EU is non-linear, and we estimate the model parameters from this non-linear structure rather than from a reduced form linear regression model. Each model parameter, however, can be made a linear function of task or participant characteristics. For example, the CRRA utility function in equation 4 has only one parameter that is being estimated, *r*, and we can make

[6] Alternatively one could use a logit model for the stochastic specification of the model, as done in Gao, Frejinger, and Ben-Akiva (2010). This should produce equivalent results, and would not affect the specification of the latent decision models. Because the data is a panel we assume that the errors are clustered on the participants.

it a linear function of a set of participant characteristics, z_k, where k is an index for each characteristic such as gender and age, in the following way:

$$u(x_{ij}) = \frac{(v - wt_{ij} - T_i)^{1-r_0-\sum r_k z_k}}{1-r_0-\sum r_k z_k} \tag{4'}$$

The utility function (4') is then incorporated in the expression for EU *Eq. (5)* and *Eq. (8)* is estimated based on the argument in *(9)*. This ML approach easily generalizes to other utility functions and to RDU.

We also estimate equation *8* assuming that φ is a linear function of p, the probability that there will be no congestion on the route, and the net outcome $(v\text{-}wt_{ij}\text{-}T_i)$, plus a vector of participant characteristics, z_k.

$$\varphi = \beta_0 + \beta_p p + \beta_u\left(v - wt_{ij} - T_i\right) + \sum_k z_k \tag{9'}$$

The argument φ in equation *8* is replaced with *9'*:

$$\phi(\varphi) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\beta_0+\beta_p p+\beta_u(v-wt_{ij}-T_i)+\sum_k z_k)^2} \tag{8'}$$

This specification is a standard linear probit model relating the binary route choices to these variables. We include this specification as a way of organizing the data in ways that may be more familiar to many readers than the non-linear latent decision models.

When estimating these models on the lottery data instead, in both the linear (*8'*) and non-linear (*eqs. 8* and *9*) form the only difference is the value of the outcomes, x_{ij}. The argument in the utility function is simply the lottery prize (y_{ij}) instead of the net value of the simulated drive:

$$u\left(y_{ij}\right) = \frac{{y_{ij}}^{1-r}}{1-r} \tag{4''}$$

4.2 Independent variables

Table 5 lists all our exogenous variables. *Prob* is the probability assigned to the worst outcome: the low, risky prize in the lottery task and the congestion outcome on *9th Ave*. in the simulated environment. *Prize* is the highest lottery prize in the lottery task and the trip value, v, in the simulated task. *Toll* is the toll charge on *7th Ave*. in the simulated task, and it varies across both participants and drives. The value argument in the utility function in the simulated task takes the *Prize* and subtracts the *Toll* (for *7th Ave*.) and then subtracts the value of lost time (if *9th* is selected and if there is congestion).

In all models we include a covariate that 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, averaged across participants. This allows us to 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 (risk averse individuals will on average earn less), we need to use an instrumental variable technique to avoid biased coefficient estimates (See Chapter 8 in Greene, 2012). We use the exogenous experimental treatment variables (*Prob*, *Prize*, and *Toll*), the actual frequency of congestion or the actual role of the die in the lottery task, and demographic variables as instruments for cumulative earnings. The correlations between the predicted cumulative earnings based on these instrumental variables and the actual earnings are above .64 for all tasks, so the instruments are good.

4.3 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 simulated driving 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%. We also see that the proportion of risky choices is somewhat higher in the simulated tasks than in the lottery tasks. Since the lottery tasks and the simulated tasks are only similar, not identical, the somewhat higher propensity to choose the risky option in the simulated 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 decision models in order to investigate whether risk attitudes differ across students and field participants, and across lottery and simulated tasks. We first discuss our findings from the linear model shown in *eqs. 8'* and *9'*.

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 Simulation, and Student Simulation. The endogenous variable is the choice of the risky over the safe option in the lottery task or *9th Ave.* over *7th Ave.* in the driving simulated task.

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 there are fewer drivers using it. First, we see that the propensity to choose *9th Ave.* in the simulated driving task 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 Ave.*, 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 participants 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 an increase in the propensity to take the risky option across task repetitions in the lottery task, but a (weakly significant) decrease for students in the simulation (*Task*). Behavior in the driving simulation is affected by the accumulated earnings, but this is not the case in the lotteries. For the driving simulation the coefficient on *CumW* is positive and significant.

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

4.4 Expected Utility and Rank Dependent Utility

Table 9 shows the marginal effects of our four models estimated using EUT with a CRRA utility function. The top part of the table shows the estimates for the parameter *r*, and the bottom half for μ. These estimates are generated through the steps outlined in equations *8* and *9*, employing the decision model shown in equations *1-5*. For each of them we show a constant term as well as the marginal effects from various covariates. All covariates are linear effects as illustrated in equation *4'*. Table 10 shows the marginal effects of the same models under RDU, again with CRRA utility and using the power utility function of equations *6-7*. The top part of the table shows the estimates for the parameter *r*, and the bottom half for γ.[7] Below we discuss first the case of the EUT model, then the RDU and finally we compare the results of these two models against the results of the Probit models.

We see in Table 9 that, under the assumption that subjects behave according to the EUT model, there is evidence of risk aversion as suggested by the positive constant term of the equation for *r*, shown in the top part of the table. Across the four models we see a much lower risk aversion for students in the simulated task than we see in any of the other models: this is the only case where the constant term in the *r* equation is not significantly different from risk neutrality. However, as students gain more experience in the task they become more risk averse in the driving task, as can be seen by the positive coefficient on *Task*. The marginal effect of only one additional drive for students is 0.31, compared to the initial difference in point estimates between students and field participants of 0.62. Thus, almost all of this initial difference disappears after the second paid drive.

But this is not the only difference we see. Field participants are more risk averse than student participants in the lottery task, and student participants are more risk averse in the lottery task than the simulated task. These attitudes

[7] We also included μ in the RDU model, but it is not shown in the table.

converge with experience (shown through the coefficients on *Task*). Assuming that the learning effect is linear across task repetitions, as this specification does, we see a significant reduction in the initial risk aversion for field participants in the lottery task from 1.26 to 0.54 after two repetitions (using the marginal effect from Task of -.36, i.e., 1.26 – 2*0.36 = 0.54), and a significant increase in risk aversion for students in the driving simulation from 0.09 to 0.71 (using the marginal effect from Task of +.31, i.e., 0.09 + 2*0.31 = 0.71), after two drives. This compares to the initial risk aversion for students in lotteries and field participants in driving simulations of 0.72 and 0.71, respectively, with neither of them showing a significant *Task* effect. The highest initial risk aversion (for field participants in the lottery task) shows the strongest decrease (-.36), followed by the second highest risk aversion category at a marginal decrease of -.20 (for students in the lottery tasks with p-value of 0.10), with no significant adjustment for field participants in the simulated task and a significant increase for students in the simulated task (by +.31). We conclude that risk attitudes are converging across the models.[8]

Our estimated EUT CRRA model includes demographic covariates, but we find very few significant effects. Female students are significantly more risk averse than male students in the simulated driving task, implying that our finding of risk neutrality discussed earlier applies primarily to the male students. African-American/Black students are more risk averse than other students in the lotteries, and African-American/Black field participants are more risk averse

[8] We estimate the same models using a CARA utility function under EUT and find qualitatively the same results. Students in driving simulations are initially not risk averse and the highest degree of initial risk aversion is found for field participants in lotteries. While the task variable indicates some convergence in preferences, it is not significant in any of the CARA models. Instead we see find risk aversion in the lottery tasks decreases with the cumulative earnings, thus approaching the preferences in the simulated task. We also estimate a mean-variance utility model using probit. We find no significant coefficients on the difference in variance between the safer and the riskier option in these models. As we can see in the EUT model estimations this does not imply an absence of risk aversion. The difference in results between these models is not surprising since, in our case, the mean-variance specification is not equivalent to an EUT model because the travel time distribution is not normal. Web Appendix B shows the CARA and the mean-variance estimates. The appendix is found among the supporting documents to DBEL-WP1301 at http://dbel.robinson.gsu.edu/publications.

than other field participants in the simulations. We find no other significant demographic effects.

Figure 3 shows the fitted distributions of the estimated risk aversion coefficients for these models, pooled across tasks. The top left panel compares the risk attitudes of students across lotteries (mean 0.34) and simulations (mean 0.16). The top right panel compares the risk attitudes of field participants across lotteries (mean 0.46) and simulations (mean 0.31). There is a left tail for the simulated task that extends into the risk loving range. The two bottom panels compare students to field participants in each of the two types of tasks: lotteries to the left and simulations to the right. From the lower right panel we also see that the distribution of risk attitudes for students in the simulations have more mass in the risk loving range than does the distribution for field participants. This tail into the risk loving range is greatly diminished as participants gain experience in the simulations.

We now focus on the results for the RDU model, as shown in Table 10. This model shows utility functions with similar patterns as the EUT model: the constant terms for the *r* coefficient of the models imply that the least amount of utility concavity is found in the simulated task and the most concavity is found in the lottery task, although there is no significant difference between students and field subjects. On the other hand, the convexity of the probability weighting function, which also indicates risk aversion, is higher for field subjects than for students, although the estimates are unreliable so none is significantly different from 1. The combined effect of risk aversion as captured by the utility function and the probability weighting function jointly is to (weakly) rank the models in the same order as under EUT. Convergence , as shown by *Task*, is only significant for the lottery task. For field subjects the reduction in risk aversion is expressed through diminished concavity of utility and for students through increased concavity of the probability weighting function (indicating a movement towards optimism.) In the simulated driving tasks we see no significant adjustment across tasks.

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 participants are less inclined to choose the risky option compared to students, at least in the lotteries, consistent with the higher risk attitudes we see in the EUT and RDU models. A higher r means that the utility from the higher earnings in the risky option is discounted more strongly, thereby lowering the EU (or RDU) of the risky compared to the safe option, and consequently decreasing the likelihood of selecting the risky option as shown in Equations *(8)* and *(9)*. A higher γ (a more convex decision weight function) leads to the same qualitative effect since the weight on the utility of the higher earnings for the risky option is lower. Further, in the Probit model we also see that field participants and students converge in risk attitudes with experience, also consistent with the models here. However, the Probit model does not reveal the differences in risk attitudes across the lottery and simulated driving tasks that we see in the structural estimations. 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 simulated task. Field participants have a significantly higher risk aversion in the lottery task than in the simulated 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 the same incentive structures so that the statistical model specifications are the same, but this can be relaxed when estimating structural decision models thus giving the researcher more freedom to design the tasks to fit the policy context.

Even though field participants 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 participants forget their experiences more easily than students, even though the amount of time that passed between Session 1 and Session 2 was the same for both groups.

However, the field participants once again react to task repetition by lowering risk aversion within Session 2. In the EUT model we see again, as we did in the Probit model, that there is a significant effect from earnings accumulation in the simulated driving task (*WS2*). 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 lower risk aversion displayed initially by students in the simulation may derive from a perception that the simulation 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. The initially much higher risk aversion displayed by field participants 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, which may be the case when we give field participants a stylized lottery task like the one used here.

Finally, we look at the sensitivity parameter, μ, in the EUT specification. 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 participants are a little more sensitive to the EU differences as the cumulative earnings increase in the simulated driving task.

5. CONCLUSIONS

In this paper we estimate risk attitudes using a methodology from Experimental Economics that marries the control of the context from Stated Choice methods with actual consequences as in Revealed Preferences. Since the consequences imposed can become expensive, we test if elicitations using less expensive participant pools and elicitation tasks lead to similar inferences regarding risk attitudes. We compare relatively inexpensive student participants to more expensive participants from the adult driving population. We also compare simple context-free elicitation tasks like lottery choices to real-time contextual tasks implemented in computer simulated driving environments.

We find that students initially behave very differently from field participants in the simulations, 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 and earnings, they will act more consistently with field participants. Thus, our findings are optimistic about being able to use students in driving simulations 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 participants, conditional on demographics. While we see that field participants are initially much more risk averse in the lottery tasks than in the simulated tasks, experience with the lotteries result in a fast convergence towards the risk attitudes revealed in the simulation task. Thus, our findings are also optimistic about being able to reveal risk attitudes through stylized lottery tasks that are relevant for simulated driving tasks. Given the high degree of variability in estimated valuations of travel time reliability found in the transportation literature, this convergence makes us optimistic about the value of using relatively inexpensive incentivized tasks from Experimental Economics for eliciting robust risk attitudes. However, attention to the presentation format is also required, as demonstrated by the initial differences in behavior.

Using students instead of field participants saves costs in recruiting, participation compensation, and staffing. Adjusting our recruitment cost to the fact that our unusually strict privacy protocol lead to a response rate of only 0.6%, and assuming instead a response rate of something more reasonable like 2%, results in a mailer cost per field participant of $22. In comparison, recruitment costs at most universities that conduct behavioral studies on a regular basis is virtually zero. Task earnings are the same for both students and field participants, but the participation compensation differs by $18 per participants. We paid each field participant $25 per meeting but each student only $7. There is also a slightly higher staffing expense in the field due to additional transportation and setup time, compared to using an existing lab at a university. We estimate this to at least $1.50 per participant. Thus, we find a $41.50 additional cost per field participant, over that of a student participant.

It is important to recognize the limitations of these findings. The payoff structure in these tasks was designed partially to be relevant for congestion pricing, but also to comply with the guiding principles for payoff structures in experiments from induced value theory (Smith 1976). While the level of the tolls are commensurate with those participants may encounter in the field,[9] their relation to the value of travel time are determined by the need to identify a range of possible risk attitudes. The induced value of total time lost in a simulated drive varies from $3.75 to $5.75, and the tolls range from $0.50 to $5.50, so that, in percentage terms, the road price that they pay to avoid this cost ranges from 9% to 96%. When the road price is a very small percentage of the value of time lost in congestion, only highly risk accepting participants would choose the risky route. On the other hand, when the road price is a very high percentage of the value of time lost, only extremely risk averse drivers would still choose

[9] For example, the range of tolls in the driving simulation was $0.50 - $5.50. Compare this to the rates on 95 Express in Miami where the minimum toll rate is $0.50 and the maximum allowable is $10.50, with peak-hour tolls as high as $7.00. Typical rates are posted here: http://173.193.186.127/pages/usage-guidelines/tolling-typical-toll-amount/southbound-95-express-typical-toll-rates. A maximum rate of $10.50 was reported by the Sun-Sentinel in February 2014 (http://articles.sun-sentinel.com/2014-02-27/news/fl-95-express-tolls-hike-20140227_1_express-lanes-current-7-mile-stretch-alicia-torrez). This article also reports that peak congestion charges are as high as $7. Based on the difference in speeds that was reported before and after the original opening of the HOT lanes the time savings for these 7 miles would be about 7 minutes, thus implying a willingness to pay of $1 per minute or more, for those who pay for HOT lane access. Assuming instead the more typical peak-rate of $4 the implied value of time would be $0.57. However, the lane choices do not reveal the full range of value of travel time Similarly, the I85 HOT lane in Atlanta can charge anything from $0.10-$0.90 per mile, for up to 16 miles Minimum and maximum rates, as well as average daily rates, can be found here http://www.peachpass.com/faq/.

the tolled, safe route. 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 see the three data collection approaches, Stated Choice (SC), Revealed Preferences (RP), and Experimental Economics (EE) as complementary. SC provides a low cost way of eliciting responses on a large set of contexts that vary in travel times and monetary costs. Unless SC is implemented using familiar references, as in Hensher (2010), or including certainty scales and cheap talk scripts as in Fifer, Rose and Greaves (2014), this approach can lead to large response biases due to the hypothetical nature of the consequences. RP, on the other hand, has real consequences but is limited in the number of contexts that can be investigated to what is already present in the field. Further, elicitation of risk attitudes using RP may be seriously confounded by variations in the respondents' perceptions of travel times that are difficult to control for. EE can provide estimates of risk attitudes, free from the possible confounds of travel time perception biases and free of hypothetical bias due to lack of incentives. The major drawback of EE is that it can become expensive, an issue we have approached in this paper.

While the evidence of the value of familiar referents in SC are encouraging, this evidence is based on comparisons of the observed choices or estimated value of travel time. It is an open question what the effects from familiar references are on estimates of decision factors such as risk attitudes and perceptions. It is possible that the lack of significant hypothetical bias in the presence of familiar references masks offsetting biases in these factors. Harrison, Johnson, and Rutstrom (2014) provide evidence that experience and familiarity with choice situations can lead to such strong prior beliefs that new information about changing circumstances, as those presented by experimental manipulations, do not affect beliefs in the way intended, thus leading to biased expectations. Brownstone and Small (2005) also suggest that respondents may ignore manipulations which they perceive as unrealistic, whether or not they really are, which is more likely when their prior beliefs are held strongly.

We argue that adding a limited Experimental Economics exercise to a policy evaluation or a planning exercise, even with a low cost student participant pool and non-contextual tasks, can generate data that is complementary to SC and RP exercises. This proposal extends the methodology of Small, Winston and Yan (2005) that used a combination of SC and RP methods along the same lines. The EE exercise does not have to be limited to elicitation of risk preferences, but can be extended to include the same behavioral measures as an SC exercise.

We agree with Fifer, Rose and Greaves (2014; p. 176) who say:

> "The practical implications of these findings provide a cautionary warning to researchers and practitioners in this field who use the results from SC models to aid in making important market decisions. Outcomes from this research will hopefully not detract from the use of SC methods but rather encourage researchers to be aware that a certain level of bias will likely be present in SC surveys and they should therefore take the necessary precautions to limit the extent of this bias. These precautions include the use of cheap talk scripts and certainty scales as well as the careful monitoring of sample characteristics (e.g. experience and demographics)."

We would add to this the inclusion of some scenarios based on familiar referents, as in Hensher (2010), and a small incentivized EE component that can be used to statistically calibrate the non-incentivized SC responses. By including both scenarios with and without familiar referents it may be possible to gain insight both into short-run responses that anchor strongly on recent experiences, and long-run responses where beliefs may adjust. The EE component will provide information regarding any bias due to lack of incentives, and can be particularly important when the aim is to not only observe but also explain choices using factors such as value of time, risk attitudes, and perceptions. By understanding choices in terms of these decision factors, it is possible to predict responses to scenarios not directly presented to the respondents.

Figure 1: Screen Shot For Lottery Practice Task

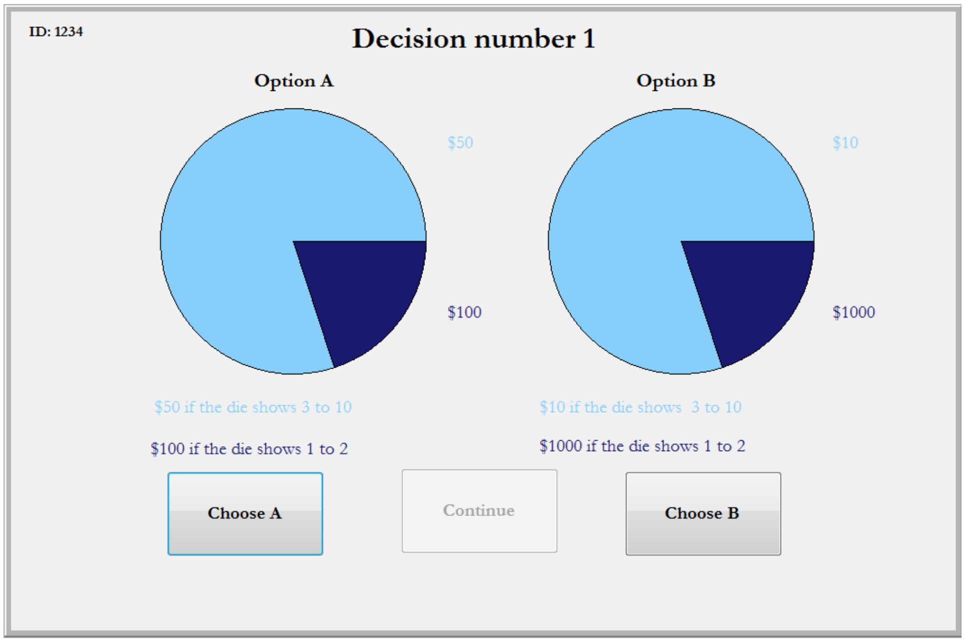

Figure 2: Downtown Network With Bus on 9th Ave.

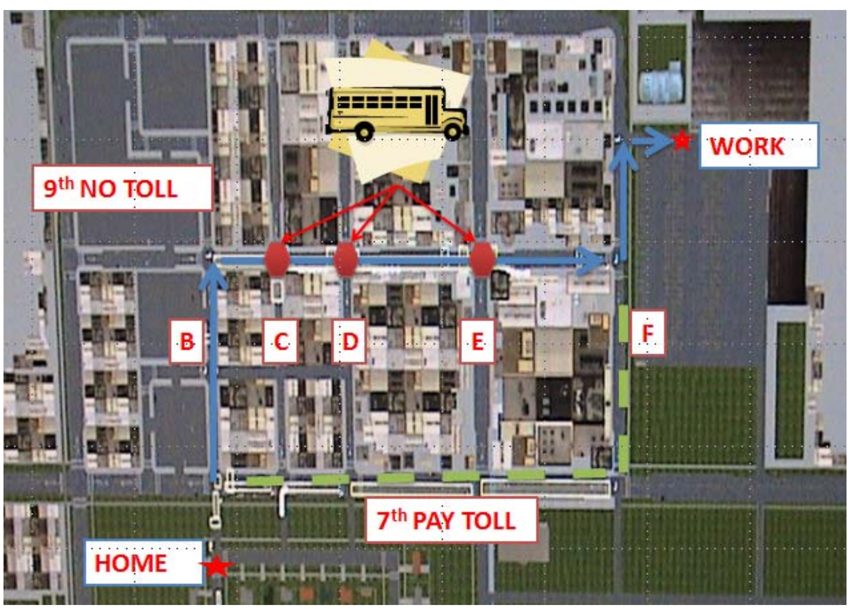

Figure 3: Estimated Distributions of Risk Aversion Coefficients for the EUT CRRA Models

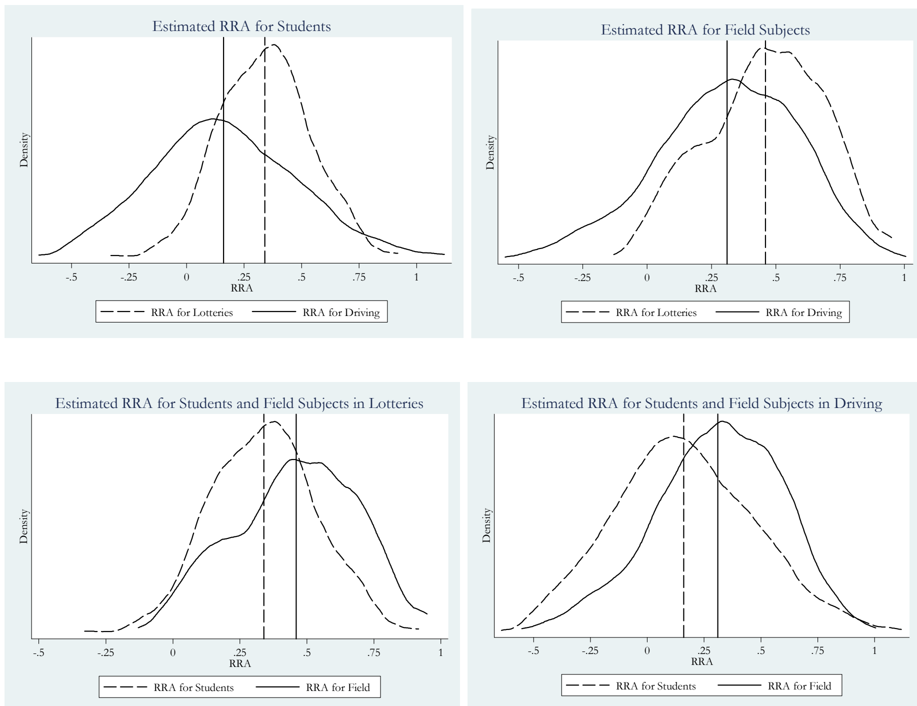

Table 1: Prizes and Probabilities in Lottery Task

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 Trip Values in the Simulation Task

	Low Toll Range	Medium Toll Range	High Toll Range
Trip Value=$4	$0.5-$1.50	$1.60-$2.50	$2.60-$3.50
Trip Value =$5	$0.5-$1.80	$1.90-$3.20	$3.30-$4.50
Trip Value =$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**				

In this example the trip value is $5, implying a delay cost of $4.75, and the illustrative tolls are $1.20, $1.90 and $3.70, respectively for each of three drives.

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: Independent variables

Variable	Description	Min value	Max value
Prob	Probability of bad outcome in risky option (congestion probability in driving simulation)	0	1
Prize	High prize in risky option	$4	$10
Toll	Toll on 7th Ave. in driving simulation	$0.50	$5.50
Task	Counter for task number, 1-4 for lotteries and 1-3 for driving simulation	1	4
Lott2	Indicator for lotteries in Session 2	0	1
CumW	Cumulative earnings[a]	0	60
WS2	Instrumented cumulative earnings (CumW) interacted with Session 2 dummy (Lott2)	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 (Oftenplay) interacted with being older than 30 years (1 - Age18-30)	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 are pooled	0	1

a) The effect from cumulative earnings are estimated using an instrumental variables approach described in section 4.2.

Table 6: Cumulative earnings by task, averaged across participants

	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 lottery tasks 3 and 4 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 Participants	**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 of Selecting Risky Option

	Field lottery	Student lottery	Field simulation	Student simulation
Constant	.18 (.483)	.88 (.012)	.29 (.486)	.35 (.417)
Prize	.08 (.002)	.08 (.004)	-.41 (<.001)	-.14 (.033)
Prob	-2.49 (<.001)	-3.75 (<.001)	-1.54 (<.001)	-1.35 (<.001)
Toll			.77 (<.001)	.31 (<.001)
CumW	-.01 (.773)	0.04 (.331)	.06 (.010)	.08 (<.001)
Task	.20 (.058)	.22 (.073)	-.05 (.554)	-.16 (.080)
Lott2	-.46 (.099)	-.32 (.323)		
WS2	.02 (.677)	-.04 (.356)		
Oftenplay	-.01 (.898)	.03 (.570)	.11 (.106)	-.03 (0.522)
Oldage_play	-.005 (.917)	-.12 (.494)	-.11 (.058)	-.14 (.75)
Female	.09 (.316)	-.21 (.078)	-.01 (.940)	-.27 (.007)
Low Income	.00 (.999)	-.03 (.869)	-.18 (.238)	.02 (.857)
High Income	-.10 (.380)	-.19 (.252)	-.04 (.765)	-.05 (.724)
African-American/Black	.10 (.370)	-.12 (.288)	-.07 (.640)	-.05 (.600)
N	1052	813	832	623
Wald Chi2 statistic	174.34 (<.001)	215.80 (<.001)	190.14 (<.001)	111.43 (<.001)
Pseudo R2	.1704	.2883	.3030	.114

The model is shown in equations 8 and 9'. Numbers in brackets are *p*-values. Coefficients are z-scores. Residual errors are clustered on the individual participant.

Table 9: EUT models with CRRA utility

	Field lottery	Student lottery	Field driving	Student driving
r				
Constant	1.26 (<.0001)	.72 (.004)	.71 (.024)	.09 (.781)
Lott2	.61 (.052)	.117 (.679)		
CumW	.02 (.76)	-.008 (.824)	-.06 (.017)	-.08 (.004)
WS2	-.03 (.731)	.014 (.696)		
Oftenplay	-.03 (.788)	-.07 (.104)	-.02 (.861)	-.03 (.659)
Oldage_play	.02 (.899)	.14 (.349)	-.009 (.949)	.06 (.551)
Task	-.36 (.002)	-.197 (.101)	.08 (.343)	.31 (.007)
Female	.02 (.861)	.11 (.305)	.11 (.393)	.34 (.035)
Low Income	.03 (.808)	-.08 (.654)	.10 (.546)	-.06 (.728)
High Income	.03 (.795)	.05 (.793)	.10 (.501)	.12 (.448)
College	-.16 (.335)		-.09 (.578)	
Age 18-30	-.08 (.813)		-.40 (.293)	
Age 56-75	-.15 (.414)		-.03 (.874)	
African-American / Black	-.12 (.381)	.14 (.004)	.19 (.024)	.16 (.275)
µ				
Constant	1.21 (.031)	1.13 (.005)	0.74 (.369)	1.71 (.023)
Lott2	-.11 (.825)	-.69 (.006)		
CumW	.05 (.625)	-.08 (.205)	0.05 (.047)	.01 (.827)
WS2	-.05 (.649)	.10 (.176)		
Oftenplay	-.09 (.687)	-.08 (.425)	-0.13 (.287)	.03 (.852)
Oldage_play	.19 (.489)	.11 (.412)	.08 (.588)	-.25 (.200)
Task	-.14 (.525)	.37 (.103)	-.02 (.855)	-.001 (..995)
Female	.42 (.220)	-.13 (.493)	.28 (.079)	-.65 (.119)
Low Income	.18 (.613)	-.42 (.122)	.14 (.430)	-.39 (.321)
High Income	-.19 (.423)	-.43 (.130)	.33 (.152)	-.20 (.629)
College	.14 (.622)		-.19 (.245)	
Age 18-30	.11 (.845)		.44 (.395)	
Age 56-75	.08 (.771)		-.03 (.875)	
African-American / Black	-.03 (.906)	-.01 (.956)	.003 (.981)	-.23 (.499)
Nsubj	253	189	254	190
Nobs	1050	811	834	623
Wald Chi2	16.04 (0.247)	10.94 (.363)	14.24 (.2201)	18.25 (.0194)
Log Pseudolikelihood	-571.46	-370.66	-407.29	-284.25

Numbers in brackets are *p*-values showing difference from zero. Further, the terms in mu are not significantly different from 1 for any of the covariate.

Table 10: RDU models with CRRA utility

	Field lottery	Student lottery	Field driving	Student driving
r				
Constant	.64 (.005)	.68 (<.001)	.14 (.619)	.25 (.437)
Lott2	.17 (.248)	-.21 (.116)		
CumW	.02 (.597)	-.02 (.393)	.002 (.725)	-.01 (.483)
WS2	-.02 (.606)	.02 (.249)		
Oftenplay	-.03 (.718)	-.04 (.053)	-.03 (.630)	-.00 (.998)
Oldage_play	.04 (.594)	.07 (.110)	.02 (.703)	-.02 (.469)
Task	-.12 (.025)	.01 (.804)	.01 (.647)	.04 (.499)
Female	.07 (.305)	.01 (.872)	.07 (.451)	-.03 (.515)
Low Income	.06 (.370)	-.15 (.058)	.02 (.682)	-.04 (.515)
High Income	-.03 (.630)	-.12 (.183)	.08 (.527)	-.01 (.917)
College	-.03 (.684)		-.06 (.497)	
Age 18-30	-.002 (.988)		.04 (.745)	
African-American / Black	-.03 (.005)	.05 (.221)	.02 (.607)	-.01 (.803)
γ				
Constant	1.70 (.006)[a]	.91 (.004)[b]	1.63 (.025)d	.92 (.004)[c]
Lott2	.54 (.221)	.38 (.070)		
CumW	.002 (.984)	.02 (.459)	-.07 (.152)	-.05 (0.101)
WS2	-0.002 (0.979)	-.02 (.466)		
Oftenplay	.02 (.896)	-.01 (.756)	-.01 (.956)	-.001 (.992)
Oldage_play	-.04 (.797)	.03 (.859)	.03 (.864)	.11 (.538)
Task	-.30 (.129)	-.20 (.027)	.10 (.497)	.15 (.297)
Female	-.06 (.663)	.12 (.210)	.03 (.879)	.26 (.107)
Low Income	.04 (.803)	.06 (.615)	-.01 (.962)	.01 (.927)
High Income	.07 (.691)	.16 (.185)	-.08 (.674)	.12 (.443)
College	-.16 (.440)		-.14 (.709)	
Age 18-30	-.12 (.738)		-.16 (.605)	
African-American / Blackblack	-.14 (.381)	.07 (.340)	.18 (.497)	.09 (.419)
μ	-.20 (.727)	-.94 (.068)	.77 (.728)	.58 (.717)
N	1050	811	834	623
Wald Chi2	10.80 (.55)	15.54 (.11)	0.96 (.9999)	0.86 (.9990)
Log Pseudolikelihood	-567.12	-364.79	-405.19	-286.53

Numbers in brackets are *p*-values. a) For Field Lottery *p*-value for difference of γ from 1 = 0.25, b) For Student Lottery *p*-value for difference of γ from 1 = .79, c) For Field Driving *p*-value for difference of γ from 1 = 0.39, d) For Student Driving *p*-value for difference of μ from 1=.80.

REFERENCES.

Andersen, S., Dixit, V. V., Harrison, G. W., Harb, R. C., Lau, M. I., Radwan, E., Rutström, E. E., Tarr. R., 2011. Experiments on Driving under Uncertain Congestion Conditions and the Effects on Traffic Networks from Congestion Initiatives. Report to U.S. Federal Highway Administration, DBEL-WP1101 http://dbel.robinson.gsu.edu/research-projects-fhwa/

Andersen, S., Harrison, G.W., Lau, M. I., Rutström, E. E., 2008. Eliciting Risk and Time Preferences. Econometrica 76(3): 583-618.

Andersen, S., Harrison, G.W., Lau, M. I., Rutström, E. E., 2010. Preference heterogeneity in experiments: Comparing the field and laboratory. Journal of Economic Behavior and Organization 73(2): 209-224.

Arrow, K. J., 1965. Aspects of the Theory of Risk Bearing. Hensinki: Yrjo Jahnssonin Saatio. Reprinted in: Essays in the Theory of Risk Bearing, Markham Publ. Co., Chicago, 1971, 90-109.

Asensio, J., Matas, A., 2008. Commuters' valuation of travel time variability. Transportation Research Part E: Logistics and Transportation Review 44, 1074–1085.

Attema, A. E., Brouwer, W. B. F., L'Haridon, O., 2013. Prospect theory in the health domain: A quantitative assessment. Journal of Health Economics, 32, 1057-1065.

Avineri, E., 2006. The effect of reference point on stochastic network equilibrium. Transportation Science, 40(4), 409-420.

Bates, J., Polak, J., Jones, P., and A. Cook, 2001. The valuation of reliability for personal travel. Transporation Research Part E, 37 (2-3), 191-229.

Becker G. M., DeGroot, M. H., Marschak, J., 1964. Measuring utility by a single-response sequential method. Behavioural Science (3): 226–32.

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

Bhat, C., Sardesai, R., 2006. The impact of stop-making and travel time reliability on commute mode choice. Transportation Research Part B: Methodological 40, 709–730.

Brownstone, D., Small, K. A., 2005. Valuing time and reliability: assessing the evidence from road pricing demonstrations. Transportation Research Part A 39, 279–293

Carrion, C., Levinson, D., 2012. Value of travel time reliability: A review of current evidence. Transportation Research Part A: Policy and Practice 46(4), 720-741.

Chaiken, S., Yaacov T., 1999. Dual-Process Theories in Social Psychology. New York, NY: Guilford Press.

Chateauneuf, A. and Cohen, M. 1994. Risk seeking with diminishing marginal utility in a nonexpected utility model. Journal of Risk and Uncertainty, 9, 77-91.

Cooper, D. J., Kagel, J. H., Lo, W., Gu, Q. L., 1999. Gaming against Managers in Incentive Systems: Experimental Results with Chinese Students and Chinese Managers. American Economic Review 89 (4): 781-804.

Cox, J. C., Vjollca, S., Schmidt, U., 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.

Cummings, R. G.; Harrison, G. W., Rutström, E. E., 1995. Homegrown Values and Hypothetical Surveys: Is the Dichotomous Choice Approach Incentive Compatible? American Economic Review, 85(1), March 1995, 260-266.

Davis-Stober, C. P. and Brown, N., 2013. Evaluating decision making "type" under p-additive utility representations. Journal of Mathematical Psychology, 57, 320-328.

Deck, C., Lee, J., Reyes, J., Rosen, C. C., 2013. A Failed Attempt to Explain Within Subject Variation in Risk Taking Behavior Using Domain Specific Risk Attitudes, Journal of Economic Behavior and Organization, March, 87, 1-24.

Deck, C., Lee, J., Reyes, J., Rosen, C. C., 2014. Investing Versus Gambling: Experimental Evidence of Multi-domain Risk Attitudes. Applied Economics Letters, January, 21(1): 19-23.

De Jong, G., and M.C.J. Bliemer (2015) On including travel time reliability of road traffic in appraisal. Transportation Research Part A, Vol. 73, pp. 80-95.

Devarasetty, P. C., Burris, M., Shaw, W. D., 2012. The value of travel time and reliability-evidence from a stated preference survey and actual usage. Transportation Research Part A, 46, 1227–1240

Dixit, V. V., Harrison, G. W., Rutström, E. E., 2014. Estimating the subjective risks of driving simulator accidents. Accident Analysis and Prevention 62 (2014) 63–78.

Evans, J., Frankish, K., 2009. In Two Minds: Dual Processes and Beyond. Oxford: Oxford University Press.

Falk, A., Meier, S., Zehnder, C., 2013. Do Lab Experiments Misrepresent Social Preferences? The Case of Self-selected Student Samples. Journal of the European Economic Association 11(4): 839-852.

Fifer, S., Rose, J.M. and Greaves, S. (2014) Hypothetical bias in stated choice experiments: is it a problem? And if so, how do we deal with it? Transportation Research Part A, Vol. 61, pp. 164-177.

Fiore, S. M., Harrison, G. W., Hughes, C. E., Rutström, E. E., 2009. Virtual Experiments and Environmental Policy. Journal of Environmental and Economic Management 57(1): 65-86.

Fosgerau, M., Karlstrom, A., 2010. The value of reliability. Transportation Research Part B 44, 38–49.

Fox, C. R., Tversky, A., 1995. Ambiguity Aversion and Comparative Ignorance. Quarterly Journal of Economics 110 (3): 585-603.

Gao, S., Frejinger, E., Ben-Akiva, M., 2010. Adaptive Route Choices in Risky Traffic Networks: A Prospect Theory Approach. Transportation Research Part C, 18, p. 727-740.

Gaudecker, H. M. V., Soest, A. H. O., Wengstrom, E., 2011. Heterogeneity in risky choice behavior in a broad population. American Economic Review 101(2): 664-694.

Ghosh, A., 2001. Valuing time and reliability: commuters' mode choice from a real time congestion pricing experiment. PhD thesis, University of California, Irvine (USA).

Greene, W., 2012. Econometric Analysis. Seventh Edition, Prentice Hall.

Harrison, G.W., Johnson, J., and Rutström, E.E., 2015. Risk Perceptions in the Virtual Wilderness, CEAR Working Paper 2015-03, Georgia State University.

Harrison, G. W., Rutström, E. E., 2008. Estimating Risk Aversion in the Laboratory. In Risk Aversion in Experiments, Research in Experimental Economics 12: 41-196.

Harrison, G. W., and Rutström, E. E., 2009. Expected Utility Theory and Prospect Theory: One Wedding and a Decent Funeral, Experimental Economics, 12(2), 133-158.

Harrison, G. W., Swarthout, T. J., 2012. Experimental Protocols and the Bipolar Behaviorist. CEAR Working Paper 2012-01, Georgia State University.

Hensher, D.A., 2010. Hypothetical bias, choice experiments and willingness to pay. Transportation Research Part B, Vol. 44(6), pp. 735-752.

Hensher, D. A., Greene, W. H., Li, Z., 2011. Embedding risk attitude and decision weights in non-linear logit to accommodate time variability in the value of expected travel savings. Transportation Research Part B, 45, 954-972.

Hey, J. D., Orme, C., 1994. Investigating Generalizations of Expected Utility Theory Using Experimental Data. Econometrica 62(6): 1291-1326.

Hollander, Y. 2006. Direct Versus Indirect Models for the Effects of Unreliability. Transportation Research Part A, 40(9), 699-711.

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

Huck, S., Müller, W., 2012. Allais for All: Revisiting the Paradox in a Large Representative Sample. Journal of Risk and Uncertainty 44: 261-293.

Jackson, W.B. and Jucker, J.V., 1982, An empirical study of travel time variability and travel choice behavior. Transportation Science 16(6), 460-475.

Kahneman, D., Tversky, A., 1979. Prospect Theory: An Analysis of Decision Under Risk. Econometrica XLVII (1979): 263–291.

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

Lapperent, M. D., 2010. Attitude toward risk of time loss in travel activity and air route choices. Journal of Intelligent Transportation Systems, 14, 166-178.

Li, Z., Hensher, D. A., Rose, J. M., 2010. Willingness to pay for travel time reliability in passenger transport: A review and some new empirical evidence. Transportation Research Part E, 46, 384-403.

Li, Z., Hensher, D. A. 2011. Prospect Theoretic Contributions in Understanding Traveller Behaviour: A Review and Some Comments. Transport Reviews, 31(1),97-115.

List, J., Lucking-Reiley, D., 2000. Demand Reduction in Multiunit Auctions: Evidence from a Sportscard Field Experiment. American Economic Review, 90(4), 961-972

Louviere, J.J., Hensher, D.A., Swait, J., 2000. Stated Choice Methods: Analysis and Applications in Marketing. Transportation and Environmental Valuation, Canbridge University Press, Cambridge.

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

Pratt, J.W., 1964. Risk Aversion in the Small and in the Large. Econometrica, 32(1-2), 122-136.

Quiggin, J., 1982. A Theory of Anticipated Utility. Journal of Economic Behavior and Organization, 3(4), p. 323-343.

Rasouli, S. and H. J. P. Timmermans, 2014. Effects of travel time delay on multi-faceted activity scheduling under space-time constraints: a simulation study. Travel Behaviour and Society, 1(1), 31-35

Reynaud, A., Couture, S., 2012. Stability of Risk Preference Measures: Results From a Field Experiment on French Farmers. Theory and Decision, August, 73(2): 203-221.

Sargent, T. J., 1987. Macroeconomic Theory 2nd edition, Academic Press Inc.

Senna, L. A., 1994. The influence of travel time variability on the value of time. Transportation 21: 203-228.

Small, K.A, Noland, R.B., Chu, X., Lewis, D., 1999. Valuation of travel-time savings and predictability in congested conditions for highway user-cost estimation. NCHRP Report 431, Transportation Research Board, National Research Council.

Small, K. A., Winston, C., Yan, J., 2005. Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability. Econometrica, v73, no. 4, July, pp 1367-1382

Smith, Vernon L., 1976. Experimental Economics: Induced Value Theory. American Economic Review, 66(2), Papers and Proceedings of the Eighty-eighth Annual Meeting of the American Economic Association, 274-279.

Stanovich, K.E., West, R. F., 2000. Individual Differences in Reasoning: Implications for the Rationality Debate? Behavioral and Brain Sciences 23: 645-726.

Stott, H., 2006 . Cumulative prospect theory. Functional menagerie. Journal of Risk and Uncertainty, 32, 101-130.

Tversky, A., Kahneman, D., 1992. Advances in Prospect Theory: Cumulative Representation of Uncertainty. Journal of Risk and Uncertainty, 5, 297-323.

Wakker, P.P., 2010. Prospect Theory for Risk and Ambiguity. Cambridge University Press, New York.

Wilcox, N., 2011. Stochastically More Risk Averse: A Contextual Theory of Stochastic Discrete Choice Under Risk. Journal of Econometrics, 1(162), May 2011, pp. 89-104.

Wilcox, N., 2015a. A Comparison of Three Probabilistic Models of Binary Discrete Choice Under Risk. Unpublished Manuscript, Chapman University.

Wilcox, N., 2015b. Unusual Estimates of Probability Weighting Functions. Unpublished Manuscript, Chapman University.

Yan, J., 2002. Heterogeneity in motorists' preferences for travel time and time reliability: empirical finding from multiple survey data sets and its policy implications. PhD thesis, University of California, Irvine (USA).