HANDBOOK FOR CONDUCTING AUGMENTED RANDOMIZED CONTROLLED TRIALS FOR EVIDENCE BASED ANALYSIS OF CONGESTION PRICING PROPOSALS

by
Elisabet E. Rutström with
Vinayak V. Dixit, Rami C. Harb, Glenn W. Harrison, and A. Essam Radwan

DBEL-WP 1401

Rutström is Professor of Economics and Director of Dean's Behavioral Economics Laboratory, J. Mack Robinson College of Business and Andrew Young School of Policy Studies, Georgia State University, USA. Dixit is Senior Lecturer, Research Centre for Integrated Transport Innovation (rCITI), School of Civil and Environmental Engineering, University of New South Wales, Australia, Harb is Senior Tolls Consultant, Atkins, NA, Harrison is Professor of Risk Management and Insurance and Director of the Center for Economic Analysis of Risk, J. Mack Robinson College of Business,Georgia State University, USA, Radwan is Professor of Civil Engineering and Executive Director, Center for Advanced Transportation Systems Simulation, College of Engineering and Computer Science, University of Central Florida, USA.

This guidebook was developed as part of a study 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.

CONTENTS

1. Introduction

a. Congestion problems and background policy issues

Traffic congestion in urban areas in the United States has become a major burden for motorists as well as government agencies. Increased trip time, excess fuel consumption, deterioration in quality of air, and rising frustration among road users are just few side effects of this growing phenomenon.

In 2011 alone traffic congestion caused 5.5 billion hours of extra time, 2.9 billion gallons of wasted fuel, $121 billion of delay and fuel cost, 56 billion pounds of additional carbon dioxide (CO_2) greenhouse gas released into the atmosphere during urban congested conditions, and 22% ($27 billion) of the delay cost was the effect of congestion on truck operations; this does not include any value for the goods being transported in the trucks (Schrank and Lomax (2012)). The cost to the average commuter was $818 in 2011 compared to an inflation-adjusted $342 in 1982. The highest total annual delay per driver is reported to be 70 hours in Chicago, Illinois. Across the ten most congested cities in the US, the cumulative delay is in excess of 200 million hours (Fleischfresser (2011)). This lost time is a financial burden to society in terms of lost wages, added cost for commodities bared by consumers, health effects, fatigue, and long term impact on lost productivity.

The federal and state departments of transportation are facing rising demand to increase the transportation network capacity. The transportation trust fund is limited and new financing schemes must be explored. Without additional revenue sources to sustain a reasonable transportation service the economy will continue to suffer. The push towards greater fuel efficiency further adds to the cost. The traditional way of funding road maintenance through federal gasoline taxes is supplying less income per mile driven. Just to illustrate this point, the proposed standard of 35.5 miles per gallon proposed for the year 2016 would reduce federal gas tax income to just over $0.005/mile. With new freeway lane-miles costing $10 million or more in urban areas (FHWA (2008)), governments are less able to “merely” add capacity each year. This emphasizes the urgency to examine increased utilization of existing capacity.

Congestion pricing provides an alternative way of relieving congestion and creating additional revenues. It has been applied in major metropolitan areas and successfully reduced congestion and improved throughput on highways. The revenue generated is then used to enhance the level of service of the facility, to provide funding for maintenance, and to possibly cross subsidize public transportation to increase the system passenger throughput. The issue of equity and possible use of generated revenue to cross subsidize transit has been widely discussed in the US and Europe provides many examples of the acceptability of this concept.

It is clear that understanding the responses to and effects of congestion pricing is important. Much of the existing evaluations of alternative pricing schemes have been done based on survey responses. A recent review by Li and Hensher (2012) reports that most studies find a low acceptability of road pricing, but that perceptions of effectiveness are important factors behind acceptability. Apart from these survey studies, evidence also exist from large scale field implementations, such as in London, England, Stockholm, Sweden, Singapore, 95Express in Miami Florida, and US-91 in San Diego California, just to name few. Such field implementations are, however, very costly ways of testing the viability of congestion pricing. In this manual we will discuss an alternative method for studying how drivers respond to congestion pricing: The Augmented Randomized Controlled Trial (ARCT).

b. Augmented Randomized Controlled Trials:

Augmented Randomized Controlled Trials (ARCT) present an alternative to stated preference approaches and traditional field tests when evaluating congestion pricing proposals or other traffic system changes.

The basic idea is to measure characteristics of the driving population that are not all directly observable, but that are important determinants to how drivers react to congestion and road pricing. Risk attitudes, risk perceptions, and subjective value of time are three such characteristics. These characteristics can then be added to simulator planning exercises, or to limited field tests, to provide more accurate predictions, and to remove both predictive noise and biases. Dixit, Harb, Martinez and Rutström (2013) provide a back-of-the-envelope example of how the neglect of such characteristics can lead to an overestimate of the acceptance of road pricing of 40% in a case of a $4 toll for a route that allows drivers to avoid a congested route that has a fifty percent chance of a delay cost of $5.75, when the trip value is $6. These numbers illustrate the importance of considering the need for measuring these characteristics.

It is the purpose of this manual to introduce the ARCT approach and to provide guidance on how to implement it. Like stated preference approaches it can be used to study and understand responses to a large set of pricing schemes under many and varied traffic conditions. Unlike stated preferences it implements choice settings with actual consequences, not just hypothetical ones, and it can use a range of realism in the presentation of the driving contexts. In this manual we will introduce contexts that use existing road networks, or simulated road networks, or that are free of any mention of road networks. Each of these alternatives has their strengths and weaknesses, as will be explained. An alternative methodology that also relies on observing choices with actual consequences is the revealed preference approach. This approach relies on observing actual route choices in existing field situations, and is therefore limited in its ability to make observations under a wide range of circumstances. ARCT brings together the ability of the stated preference approach to allow a wider range of circumstances to be studies, with the ability of the revealed preference approach to implement real consequences.

ARCT can be used in at least two different ways: as a substitute or as a complement to large scale field tests. As a substitute to field tests it would be used to augment simulator studies to improve their predictive abilities. As a complement to field tests it would decrease the required sample sizes or increase the generalizability of the findings.

ARCT has its roots in the methodology of economics experiments. Dixit, Ortmann, Rutström and Ukkusuri (2014) review applications to transportation using this methodology. Figure 1 shows the results of a Google Scholar search for papers that fit the key words “congestion pricing” and “experimental economics”, and we can see that the number has doubled since 2009. Many of these do not use salient incentive structures, however, and thus do not fit into our ARCT (or economics experiment) definition. Apart from experiments that directly tests for the Braess Paradox or the Downs-Thompson Paradox, the studies can be grouped into those that focus on equilibrium, and those that focus on individual driver behavior. Selten, Chmura, Pitz, Kube and Schreckenberg (2007), Chmura and Pitz (2004) and Chmura, Pitz and Schreckenberg (2004) approach route choice equilibrium from a reinforcement learning perspective and identify two types of drivers: direct respondents who switch route after experiencing bad outcomes, and contrarian respondents who switch in response to good outcomes. Anderson, Holt and Reily (2008) and Denant-Boemont and Fortat (2013) approach congestion from a market entry perspective where one choice alternative has a fixed cost, but the other one has a congestion externality. Convergence to equilibrium is achieved whether the congestion cost is linear or highly non-linear. A number of papers also exist that investigate departure time equilibria in experimental settings. Ziegelmeyer, Koessler, My and Denant-Boemont (2008) conducts an experiment that varies both group size and information on history of choices but find neither having an effect on behavior. Daniel, Gisches and Rapoport (2009) find that even when there is rapid and consistent convergence to equilibrium at the macro level there is considerable variation at the individual level, suggesting that mixed strategy equilibria concepts may be good modeling tools.

Individual driver behavior is the subject of Dixit, Harrison and Rutström (2014). They estimated risk attitudes and risk perceptions in a driving simulator implementation of gap acceptance in a left turn at an intersection. They find that both risk attitudes and risk perceptions vary across participants, and choice cannot be explained by only one of these characteristics. Nielsen and Jovicic (2003) conduct a field experiment in Copenhagen, Denmark, tracking route choices using GPS devices and manipulating the prices of many routes. They estimate value of time, but do not control for risk attitude or risk perceptions. They generally find that responses to price increases are stronger than predicted, which could be explained by risk attitudes.

Thus, while there is a growing literature in transportation in general and in congestion pricing in particular that uses the ARCT methodology for evidence based research, the number of studies is still very limited.

2. Why Augmented Randomized Controlled Trials (ARCT)?

In order to generate an understanding of driver behavior that is applicable across a wide range of traffic conditions it is necessary to control for influences of spurious factors. The methodology of controlled trials, or experiments, has been a hallmark of good science for many centuries and has increasingly been applied in the social and behavioral sciences over the last 150 year. The awarding of the Nobel Prize in 2001 to Vernon Smith and Daniel Kahnemann attests to the acceptability of this methodology in the social and behavioral sciences. In this section we will discuss the importance of two necessary elements of a good controlled trial design, viz. control and randomization, and a third element, augmentation, that greatly improves the universality of the findings in the presence of behavioral heterogeneity. We will also discuss alternative approaches to how to incentivize respondents, and three levels of observational aggregation: the level of the individual, the group, or the system.

a. Control: The value of wind tunnels

To demonstrate the value of controlled experimentation we first refer to the widely known use of wind tunnels in aircraft design. Imagine that one wants to understand the performance properties of a new aircraft design. Instead of building one or several prototypes and putting them into air traffic for observation, a prototype is employed in a wind tunnel trial instead. The reason for this is rather obvious: an untested craft in the air may result in damage not only to the prototype and its pilot but also to other lives and property in case of a crash. Further, a field test may not generate all the weather conditions that could affect the performance of the craft. A wind tunnel can be used to generate a variety of weather conditions, plus allow observations on performance variables that are difficult or impossible to observe in the air. Further, any damage due to unforeseen reactions of the craft can be kept to a minimum.

Testing new transportation systems is not much different from testing new aircraft designs. Installations or implementations in the field can be very costly with risky consequences due to limited understanding of the empirical properties of the system. Take congestion pricing, for instance. The ultimate effects on the volume, density and speed of traffic depends on many factors, some that are due to the properties of the road network, and others that are due to behavioral responses among the drivers. Just like in the case of field testing a new aircraft, when testing a congestion pricing design it is costly and sometimes impossible to generate the full set of conditions that would be necessary in the field. A wind tunnel approach for testing a congestion pricing system would involve constructing a test site where the relationship between the performance of the system and each individual factor could be studies. Such test sites can be constructed in laboratory environments and do not need to involve actual roads.

In this document we will show various ways that "wind tunnel" projects can be constructed that could reveal important properties of the traffic system, including the behavior of the drivers. In order to test these properties it is necessary to first identify all factors that influence the system and then proceed to test them either one at a time, or jointly. It is the purpose of theory to identify such factors. A careful ARCT design must therefore start with an understanding of behavioral theory. As we introduce examples of study designs below, we therefore start with references to theory and allow those to guide our identification of factors that should influence the performance of the traffic system.

b. Randomization: Avoiding confounds

Even in well-designed test circumstances there will be a number of factors that influence drivers' behavior that cannot be controlled for. Humans are complex creatures and actions are guided by a complex combination of genetic material and past experiences that interact with the test site environment in many ways. One accepted way of dealing with these nuisance factors is to assume that they are randomly distributed in the target population, and to make sure that respondents are randomly allocated to the trial treatments so that there is no confounding influence on the treatment effects from such uncontrolled factors.

As an illustration of how randomization solves the issue of confounds, imagine a route choice scenario where the test is about imposing a time of day toll on one of the routes. Assume for argument's sake that men are more flexible in their departure time than women. In this case, if only women are observed on the route with a toll and only men on the route without a toll we would underestimate the effect that the toll will have since women react less to the toll. However, if male and female participants are randomly allocated to both of the routes we would get an unbiased estimate of the effect since we would expect the gender allocation of the sample to each route to match the gender allocation of the target population. Of course this illustration is very extreme, but all that is required for biased findings is that the allocation of men and women to each route is slightly different from the allocation in the population. If the suspected characteristic is as easily observed as gender then researchers could easily control for these effects by simply including gender as a covariate in the analysis, as long as there are some of both men and women on both routes. However, not all characteristics of individuals that correlate with route choice behavior are as easily observed as gender, and many of them would not even be known to correlate with route behavior. In such cases randomization is the only way to avoid confounding the effect on route choices.

c. Augmentation: Understanding heterogeneity

Many research trials in the social and behavioral sciences rely entirely on randomization to treatment to control for the numerous individual effects that could influence choices. The data analysis is then based on testing if the treatments generate a shift in aggregate behavior. This is an entirely appropriate approach, but if the results are going to be universally applicable it would require very large samples in order to guarantee that all possible uncontrolled for factors are represented in the sample. Further, it does not allow the researcher to study variation in responses across individuals and the particular causal interactions between these factors and the performance of the system. For the purposes of understanding behavior in response to congestion pricing, such response variations can be crucial in order to avoid large scale field tests on a broad population. Variations in responses are likely to depend on many characteristics of both the drivers themselves, and of the traffic system.

Augmenting the randomized controlled trial with additional observations would both decrease the required sample size and allow some inferences regarding heterogeneity in responses. One example would be to add demographic observations, such as gender, income, and age, and to analyze the responses conditional on these demographic. Of course, not all individual characteristics that are of interest can be directly observed. Some of these characteristics would have to be inferred from observed behavior. Thus, the test trial would have to be designed to include enough variations in exogenous treatment conditions to identify not only the route choice itself, but also such latent characteristics.

Several individual preference characteristics present themselves as candidates for determining route choices. First, since routes differ in travel time one important characteristic is the subjective value of time of the respondent. The value of time is simply the value of foregone opportunities while travelling. For example, a mother commuting to and from work, needing to drop her child off at daycare on the way and later pick her child up again, may worry about the consequences of arriving late either to work in the morning or to the daycare in the evening, or both. An AC repairman may worry about having to work extra hours to finish his list of clients before the day is over as a consequence of travel time delays under congested traffic conditions. A teenager on the way to the mall may be annoyed over the lost social time due to delays in traffic. These examples illustrate how the subjective value of time may vary both across individuals and across trip purposes. Second, the travel time for any specific trip is not know with certainty, even though the availability of real time traffic information makes the knowledge of travel times more precise. Another factor that will affect route choices is therefore the individual driver's expectation about the travel times on various routes. Such expectations can also change if there are changes to the traffic environment such as the introduction of toll roads or toll lanes. Under those circumstances it is not only the expectation that matters, but also how individuals adjust their expectations to the new conditions. Finally, with unreliable travel times the attitudes towards such unreliability are important individual characteristics that influence the route choices. Some individuals will be more averse than others to taking routes with a great deal of unreliability in travel times.

Since all these characteristics vary across individual drivers they are important determinants to the heterogeneity of responses to congestion pricing. Understanding them is essential to the ability of making general predictions about the effects of various congestion pricing schemes.

d. Motivation to generate truthful responses.

The cheapest and most commonly used tool to collect behavioral data is a survey. Surveys can be used to collect individual characteristics, such as age, income and gender that may correlate with route choice behavior and therefore capture some heterogeneity in choices. Many of these characteristics can be verified by researchers since they are directly observable, but others require a trust in the respondent that they are telling the truth. It is a common practice when using demographic characteristics generated through survey responses to simply assume that the respondent is telling the truth. However, it is easy to imaging questions for which this may not be the case. Say, for instance, statements about ones willingness to take on risk. In many surveys this takes the form of a scale where the respondent indicates the degree to which he or she is willing to take on risk. Such survey questions do not give the respondents an incentive to report the truth, and there is evidence that respondents frequently overstate their willingness to take on risk.[1]

Surveys have also been used to find out how drivers would respond to various transportation alternatives, including congestion pricing. In these it is the route choice itself that is the focal point of the questions, not the characteristics of the respondent. The advantage of using surveys is that the researcher can easily and cheaply generate a large number of contexts, defined over traffic conditions and road pricing, and find out what the respondent thinks he or she would do under each. Due to the lack of incentives to tell the truth, or to think very carefully about ones responses, it is likely that there are biases in stated preference responses. Apart from the evidence based on eliciting risk attitudes, referred to earlier, there is

[1] Evidence of biases in hypothetical responses have been presented in Battalio, Kagel, and Jiranyakul (1990), Holt and Laury (2002), Harrison, Johnson, McInnes, and Rutstrom (2005), and Charness and Viceisza (2011).

also direct evidence of stated preference biases in the context of eliciting a willingness to pay for certain goods (Cummings, Harrison and Rutström (1995)). Since route choice in the context of congestion pricing is an elicitation of willingness to pay, similar biases should be expected. The alternative is to generate real choice situations with actual consequences. In this manual we will discuss how this can be done.

Vernon Smith (1976) introduced induced value theory as a guide to incentivize the tasks. He introduced three conditions for a good incentivized design: 1) Monotonicity: the reward medium must be chosen so that participants prefer to have more of it rather than less; 2) Salience: the task design must be such that the relation between the choices made by the participant and the reward this results in matches how actions lead to consequences in the intended field application; and 3) Dominance: any other influences on the participants behavior must be minor in comparison to the reward medium. Commonly used incentive schemes use monetary payments.

Even if one uses incentives that fulfill these requirements however, it is essential to also worry about other details in these incentive schemes. Some payment schemes are more transparent and easy to understand than others. In addition, depending on what theory of decision making one relies on to identify the characteristics that influence choices, there will be constraints on what payment mechanisms that are consistent with the theory. For example, a common payment approach is to give each respondent several choice tasks, but to only pay for one of them, selected randomly. The advantage of such a random payment mechanism is that one can give the respondent a strong incentive in each choice task without having to pay a large amount of money. For example, if one gives a respondent 10 choice tasks, each with an expected payment of $100 (high enough to give the respondent a strong incentive to think about the choice carefully), paying for all tasks would cost the project $1,000 per subjects. This high per subject cost would likely severely limit the number of subjects that could be included within the project budget. By only paying for one out of the ten tasks, the subject pool could be 10 times larger.

Problems arise with this random payment approach for at least two reasons. First, when presented with a series of choice tasks, subjects frequently engage in what is referred to as “switching” behavior.[2] This is when subjects act as if their preferences are not stable across choices, or as if they are completely indifferent between the choice options for a wide range of tasks. There are many reasons why this could happen: respondents may simply be confused because they are asked to make choices in a task setting that they are completely unfamiliar with; or respondents may mistake the task for one where they are paid for every task at the end of making all their decisions, and they think it best to vary their responses so as not to put all their eggs in one basket. The second instance when this random payment mechanism creates problems is when one wants to assume that preferences over risk arises out of two separable evaluation processes, as suggested by Rank Dependent Utility theory and Prospect Theory.[3] REFERENCES In one process subjects vary in their sensitivity to differences in payoffs depending on the size of the payoff. For example, when comparing $10 to $100, a tenfold increase in the money value, they may subjectively only value the increase as an 8 fold increase, thus discounting the attractiveness of the higher payoff. Such relationships between payments and subjective values are captured through the properties of utility functions. In a second process subjects vary in their perception of likelihoods, or probabilities. An optimistic subject, for instance, would subjectively exaggerate the likelihood of a desirable outcome, while a pessimistic subject would do the opposite. If a researcher wants to investigate the appropriateness of this dual cognitive process it is not possible to use a random payment protocol without adjusting the empirical estimation approach to reflect the added randomness that this protocol face the subject with. This can quickly result in very large and unwieldy empirical specifications.

[2] Charness and Viceisza (2011).
[3] See Quiggin (1982, 1993) for Rank Dependent Utility Theory, and Kahneman and Tversky (1979) and Tversky and Kahneman (1992) for Prospect Theory.

Thus it is important to worry not only about how to design each task, but also about how to motivate the respondents. It is not as simple as asking whether or not to use money incentives, but considerations of how many, and which tasks to pay for are equally important.

e. Levels of aggregation of evidence

Ultimately the purpose of congestion pricing is to affect the volume, density and speed of traffic. In this sense the analytical interest lies at the system level. However, the characteristics of the decision units, i.e. the drivers, will matter to the performance of the system, unless the sample sizes are large enough to rely on randomization (as discussed earlier) to deal with confounds arising from response variations across individuals. Apart from studying the system as the unit of analysis, a transportation expert may therefore also value an improved understanding of both the individual driver, but also smaller aggregation units, such as groups of drivers, or groups consisting both of drivers and agents that the drivers interact with. One example of such groups is the group of contracting parties in shipping arrangements. Such groups would include the shipping client, perhaps a retail outlet, the shipper who arranges the logistics, and the trucker who carries out the delivery. The characteristics of relationships in such groups will also matter for the performance of the system.

A fuller understanding of all the effects from various congestion pricing schemes may therefore require analysis at the individual driver level, at the group interaction level, and at the system level. The methodology of ARCT can generate data appropriate for any of these analytical levels. This manual focuses on how to elicit characteristics at the individual driver level.

3. Designing and implementing an ARCT

In this section we will introduce the basic principles behind a good ARCT design: relying on theory to guide the design to control for confounds and identify behavioral factors that influence choices. After this introduction we will, in turn, discuss each design element in more detail as well as several important issues in implementation.

a. Basic Principles: Knowing what to measure

The purpose of theory is to identify important logical connections, i.e. to identify how behavior is determined by factors such as participant characteristics or characteristics of the travel context. By properly understanding this logic it is possible to test if certain connections are true and to apply the induced value theory properly. For example, a very simple decision theory is Expected Value (EV). The premise of this theory is that when a choice is made it depends on only two things: the likelihood of various events happening and the outcome of those events.

For example, imagine a driver making a choice over two routes, A and B, driving between an origin and a destination. Perhaps the origin is home and the destination is the airport. Imagine further that each route can have two outcomes: you arrive at the destination on time (to catch your flight) or you arrive at the destination too late (and miss your flight). Each outcome, catching your flight and missing your flight, has some value, perhaps monetary. Imagine that, if you miss your flight you have to spend money on another ticket, but if you arrive on time you do not. Thus, the monetary value of catching your flight is the full value of taking the trip (say, $1000) and the monetary value of missing your flight is the trip value net of the price of the last-minute ticket (say it is $500). As you approach the point where you have to choose between route A and B your travel time information phone app tells you that if you take route A there is a 10 percent chance of getting to the airport on time, while if you take route B there is a 70% chance. EV decision theory then states that you will take the route that has the highest probability weighted value. For route A the EV is

(1) $$EV_A = 0.1 * \$1000 + 0.9 * \$500 = \$550$$

There is a 10 percent chance of arriving on time and getting the full trip value ($1,000) and a 90 percent chance of arriving late and having to buy a last minute ticket, leaving only $500 in value for the trip. For route B the EV is

(2) $$EV_B = 0.7 * \$1000 + 0.3 * \$500 = \$850$$

There is a 70 percent chance of arriving on time and a 30 percent chance of being late. Thus route B has the highest EV since it has the smallest chance of arriving late. The propensity to choose route A is increasing in the difference between EV_A and EV_B, thus increasing in the probability attached to the higher money amount, and increasing in either of the money amounts offered on route A. If drivers really behaved in the fashion described by this theory it would be a straight forward exercise to design trials that vary in probabilities and money amounts and then estimate a linear function connecting observed route choice propensities to probabilities and consequences. Nothing else would be needed. However, while simple, this model is a poor description of how human choices are really made, and we will explore more sophisticated approaches in this guidebook. Before we do so, we will explore some other concerns that affect how good your data will be.

Assume that you have collected data on route choice over some period of time and you want to construct a prediction model that captures the effect of various congestion levels on route choice. Imagine that you estimate that the propensity to choose route B, the relatively faster route, is increasing in the travel time difference between A and B, at a rate of 0.3 per minute difference. However, what you do not know is that during the time that you collected this route choice data, the airlines also increased the price of last-minute tickets. Since in this illustration we are assuming that the correct decision model is EV, this makes a late arrival more costly to the driver and this should therefore, on its own, increase the propensity to take the route that is less likely to lead to such additional costs, i.e. the B route. Your estimate of 0.3 is therefore overestimating the reactions to changes in the travel time differences since it is confounded by the simultaneous increase in the cost of delay. By first formulating your decision theory you will understand that you have to measure not just the changes in the travel times, but also any changes in the value of later arrivals, in order to construct a reliable prediction model. This understanding will then guide your study design to observe both travel times and outcome values.

There are many competing decision theories, and EV is only one of them. A popular alternative is Expected Utility theory (EU), which assumes that as outcome values vary, the subjective valuation of these outcomes do not change in direct proportion. Thus, individuals differ in their sensitivity to outcome variations. If you were estimating the effect on the route choice propensity as a function of changes in the outcome values, it would be non-linear. This phenomenon is usually modeled through a utility function u=u(x) where x is a typical outcome. It is commonly understood that utility is increasing in x, but that the degree to which this occurs varies across individuals. If the utility function is concave, then individuals discount higher values of x relatively more than they discount lower values of x. To return to our numeric example above, the outcome of $1,000 should be discounted more than the outcome $500, so the subjective utility values may instead be 800 and 450, respectively. We can then calculate the EU of the two routes as

(3) $$EU_A = 0.1 * 800 + 0.9 * 450 = 485$$

(4) $EU_A = 0.7 * 800 + 0.3 * 450 = 695$

Not only are both of the EU values smaller than the EV values, due to the discounting, but the difference between them is also affected so we should expect that, if drivers behave as predicted by EU the route choice propensities would also be affected. This would show up as non-linear effects on choice propensities. With this decision theory in mind a well-designed study would have to observe not only the travel times and the outcome values but also the varying degrees of sensitivity across individuals.

A simple extension to estimating route choice propensities as a linear function of probabilities and outcomes is to introduce quadratic terms. This is a very restricted way, however, and may not generate very general prediction models. Following Small, Winston and Yang (2005) a more general approach would be to estimate route choice as a function of the difference in mean travel times and the difference in the variance of travel times across routes. However, the most flexible approach would be to estimate the route choice propensities as a function of the non-linear decision models directly. We will refer to this approach as structural maximum likelihood (SML) estimation.[4]

The SML approach easily allows the estimation of more sophisticated decision theories, such as Rank Dependent Utility (RDU) and Prospect Theory (PT) which add behavioral factors to those discussed here. If these are to be estimated they would require additional design modifications in order to identify these added factors. Each of these models build on each other, so by identifying all factors proposed by the more extensive theory, PT, it is also possible to test the validity of each.

The understanding that individual drivers vary in many of these behavioral factors imply that they need to be accounted for in the analysis of the route choice data, or else the estimated choice propensities will be confounded and therefore biased. In the following sections we will discuss various ways to design a study to properly account for such confounds and avoid biased estimates.

i. Controlling for confounds

1. Direct control

Directly controlling for a confound implies that one must have a way of identifying its effect. As discussed earlier, theory is our tool for understanding both the presence of a factor that influences route choices, and also how it does so. By hypothesizing how it affects decisions it is possible to isolate and identify the factor and the degree of its influence. Returning to our earlier example of driving to the airport, we argued that the price of last minute airline tickets would be a factor that affects the route choice, since it affects the value of the consequence of missing the flight. Direct control would imply including the price of late minute airline tickets in the utility functions as shown in equations (1) – (4). For the case discussed earlier where men are more flexible in their departure time than women and the purpose is to estimate responses to road price changes, direct control would imply that we include gender as a factor in the analysis.

2. Control through randomization

Direct control works well when a possible confound is known and observable. However, in every case there will also be a number of confounds that are not even known, or at least are not observable. For

[4] This is the approach used in Andersen, Harrison, Lau and Rutstrom (2008) and Harrison and Rutstrom (2008).

example, it is possible that one route has several bottle necks leading to a lot of stop-go traffic flow, and this may be a factor that influences a driver who has a strong preference for continuous flow traffic but that does not influence those who do not care about that. Or, one route may go through an industrial area that is aesthetically unappealing, and some drivers may put a value on avoiding such a route, while others do not care. In each of these cases it may be difficult to place a quantifiable value on the consequences for each participant, in which case one must rely on the assumption that a randomly recruited sample of participants is representative of the distribution of these values in the population so that, on average, the estimates of how route choices are affected by the quantifiable values (such as the trip values and the delay costs) are unbiased.

3. Bringing unobservables under direct control

Some behavioral factors may appear not to be amenable to direct control at first glance, because they are not directly observable, but methods have been developed to measure them in indirect ways in order to include them in the analysis. These values include the risk attitudes of the drivers, and the perceptions that drivers have over probabilities and outcomes, including the subjective value of time. Since all of these factors will differ across drivers, and sometimes also across traffic systems, it is important to be able to infer what they are. In Section 3b we will outline the methods for drawing such inferences, and several instruments to implement them. First, however, we will discuss how each of these factors may affect route choices.

Risk attitudes

The term risk attitude refers to when agents may have an aversion to taking on risk, when other options exist that are otherwise equal but have no, or at least less, risk. Risk in this sense may not involve losses, but it does involve how likely various outcomes are. Take equations (1) and (2), for instance. In these equations we are looking at two routes, A and B, that lead to the same outcomes, \$1,000 or \$500, but with different probabilities. One way of measuring risk in this case is through the variance of the outcomes. We can show that the variance is higher for route B than for route A:

(5) $$VAR_A = 0.1\,(1{,}000 - EV_A)^2 + 0.9(500 - EV_A)^2 = 20{,}250 + 2{,}250 = 22{,}500$$

(6) $$VAR_B = 0.7\,(1{,}000 - EV_B)^2 + 0.3(500 - EV_B)^2 = 15{,}750 + 36{,}750 = 52{,}500$$

Thus, a risk averse agent should discount the value of taking route B compared to taking route A, since B is riskier. One possible way to do this is to simply use a utility function that is concave and therefore discounts the added utility from better outcomes, as shown in equations (3) and (4). By placing lower subjective values (800 and 450) on the outcomes (1,000 and 500) we saw that the EU for route B is 155 less than the EV, but the EU for route A is only 65 less, therefore narrowing the value difference between the two route options. The extent to which better outcomes are discounted will differ across participants. In some cases it is even possible for the rank ordering of the options to change due to the concavity of the utility function. To see this, imagine two routes that have the same EV, but different route values and probabilities:

(7) $$EV_A = 0.5 * \$1{,}000 + 0.5 * \$500 = \$750 \; where \; VAR_A = 0.5\,(1{,}000 - EV_A)^2 + 0.5(500 - EV_A)^2 = 31{,}250$$

(8) $$EV_B = 0.1 * \$3{,}000 + 0.9 * \$500 = \$750 \; where \; VAR_B = 0.1\,(3{,}000 - EV_B)^2 + 0.9(500 - EV_B)^2 = 562{,}500$$

In this example we have changed both the probabilities and the value of the best outcome. For agents that strongly discount the value of the additional \$2,000 they get in option B, option A will now be more attractive even though in expected values the two routes look the same. Thus, these agents would choose route A based solely on the fact that it is less risky.

However, favoring a less risky route can also be due to how pessimistically an agent views his chances on a risky route. Pessimism would imply that an agent sees a good outcome as less likely than what the actual probabilities would indicate, and optimism would imply that an agent sees a good outcome as more likely than what the actual probabilities would indicate. In equations (7) and (8) for example, a pessimistic driver may act as if the probabilities of arriving on time are lower than the 0.5 and 0.1 that they really are. In equation (7) for example, he may feel that it is .3 instead of .5, and in equation (8) that it is .01. This pessimism changes his perception of the EV to a subjective one (SEV) for the two routes in the following way:

(9) $SEV_A = 0.3 * \$1{,}000 + 0.7 * \$500 = \$650 \; where \; VAR_A = 0.3\,(1{,}000 - EV_A)^2 + 0.7(500 - EV_A)^2 = 52{,}500$

(10) $SEV_B = 0.01 * \$3{,}000 + 0.99 * \$500 = \$525 \; where \; VAR_B = 0.01\,(3{,}000 - EV_B)^2 + 0.99(500 - EV_B)^2 = 61{,}875$

We see here that the pessimism over the likelihoods leads to the SEV being less than the EV on both routes, but more so on B than on A. Thus, pessimism leads to less demand for risky routes like B. Optimism, then, would lead to the opposite: increased demand for the risky routes.

So we have now demonstrated how at least two types of behavioral factors that are not directly observable influence route choice:

- Varying sensitivity to values implied by the consequences
- Optimism and pessimism over likelihoods

Both of these factors are considered risk attitudes, but they imply different reactions to changing traffic conditions depending on whether these changes affect the probabilities or the consequences.

Subjective beliefs

Another important factor that determines route choices under congestion risk is the beliefs that drivers have about the likelihood and level of congestion. When observing the risky route choices of drivers, it is important to not attribute their selections only to their risk attitudes. It is also likely that drivers vary in their beliefs about the likelihood of congestion. For example, Small, Winston and Yan (2005) show that drivers who take the SR91 Expresslane in LA overestimate the likelihoods of large time savings for this lane, thus demonstrating that beliefs about probabilities do not reflect what the actual probabilities are.

Variations in the beliefs over risks of travel delays is not the same as optimism and pessimism over risks that are known. Even individuals who are neither optimistic nor pessimistic can have beliefs that are biased in a positive or negative direction simply because they have limited knowledge about what the true likelihoods are. Because of this it is important to present subjects with precise information about the likelihoods of various outcomes when assessing risk attitudes, or else the possible biases in the subjective beliefs will confound the risk attitude inferences. Further, while it is likely that an individual's optimism or pessimism is something that affects all of their decisions in similar ways, biased estimates of unknown probabilities will depend on the context in important ways and can therefore not be relied on to be constant across decisions.

Apart from assessing whether drivers who are affected by proposed road pricing have beliefs that are biased, it is also important to assess how these drivers update these beliefs as they encounter new situations. It is likely that there are various types of biases in such subjective updating processes. For example, some people may exaggerate the influence of specific events on their beliefs about future recurrences.

Section 3b will show several ways for eliciting beliefs and belief updating processes from individuals, and how to control for these beliefs when inferring risk attitudes. Unlike risk attitudes, beliefs are likely to vary across driving contexts and traffic systems, both because the information that is available to drivers will differ, but also because the relationship between available information and a driver's personal experience will differ.

Value of time

A final factor that will differ across drivers is the value of time. The consequences of being delayed will vary across drivers, and values (monetary or non-monetary costs) associated with these consequences will also vary. For some individuals and circumstances these consequences are discrete and for others they vary continuously with the time delay. For example, somebody on her way to work as a receptionist in a medical clinic may face a stiff penalty, perhaps even lose her job, is she arrives late, because her delay leads to delays in the patient intake process. In such a case it is not each minute of delay that is costly to her, but the possibility of being delayed enough to arrive beyond a specific threshold time. On the other hand, a person who is on his way to the mall to do some shopping is not facing such a discrete delay cost, but may rather feel that each minute delay imposes a cost since it reduces his shopping time.

Equations (1) and (2) are examples of the former, discrete cost case. For a continuous cost case, equation (1) could be rewritten as:

(9) $$EV_A = \int_{t=0}^{T} (\$250 - \$2 * t)dt$$

which shows a person who values making the trip at \$250, but has a cost of delay that is \$2 per minute. Even when the cost of delay is not monetary, it can be argued that non-monetary costs can be evaluated in terms of money in almost all cases. In our daily lives, we all spend money so that we can avoid spending time on performing many tasks ourselves: we buy lunch instead of making it, we pay a cleaner instead of doing the cleaning ourselves, and we pay for gas to get to a destination faster than we would by foot or bike. In fact, the idea of congestion pricing relies on exactly this argument: a reduction in the travel time leads to benefits that can be compared to the amount of money that must be paid to access the faster route.

In section 3b we will show procedures for eliciting value of time from drivers who are candidates for choosing a priced route. Unlike risk attitudes, but like subjective beliefs, value of time will vary across traffic systems and trip purposes, in addition to many characteristics of the driving population.

ii. Identification and power

With so many unobservable factors that influence route selection it is necessary to worry about the ability to identify all of them. Once we have found procedures that can provide such identification, it is also necessary to worry about maximizing statistical power, i.e. the ability to reject certain values if they are false. We will discuss identification around our task designs, and refer readers to standard statistical texts for discussions on power.

Generating an ability to identify all required parameters simply involves the mathematical criterion for solving a system of equations: having as many equations as unknown variables. For example, if you want to identify risk attitudes not only through the curvature of the utility function, i.e. the sensitivity to changes in the values of the consequences, but also through optimism and pessimism, it is necessary to observe the choices of subjects in more than one risky context. Changes in contexts generate the additional equations that are necessary to identify the additional factors. We will show a number of examples of how to estimate multiple factors.

b. Task design and examples

Having discussed several important principles behind a good ARCT design we will now demonstrate several task designs that can be used to elicit the factors that determine individual driver's responses to congestion pricing. Since the risk attitude is such an important determinant to route choices, we will start by showing how tasks can be designed and carried out to get good measures of this. Recall first that risk attitudes can be expressed either in the form of decreased sensitivity to value changes as the overall value gets higher, or pessimism about the probability of something good happening. We will focus here on the former case, capturing risk attitudes that are expressed through the sensitivity to value, i.e. the concavity of the utility function.

We will start with a discussion of several tasks that have been used in the literature to elicit risk attitudes, and pay special attention to strengths and weaknesses in various approaches. These tasks vary in how the risk is being presented to subjects and in how they are being paid. Both of these aspects have consequences for the type of inferences that can be drawn. We will then give an example of what the instructions can look like, followed by an overview of tasks that can be used to elicit additional characteristics. We also discuss study protocols and give an example of a study design that uses multiple tasks to elicit multiple individual characteristics for route choice behavior.

i. Tasks for eliciting risk attitudes

Table 1 shows a set of pairwise choices with monetary outcomes, which could be presented to a participant. It is not necessary for the outcomes to be monetary, but it is convenient for our discussion here to do so. The outcomes could be framed as travel time outcomes instead, but in such a case additional tasks would be necessary in order to measure the participants' subjective value of time.

In each task in Table 1 there is an option that is completely risk free, shown in the second column, and another one that may result in a higher or lower outcome, with some probability. The higher, risky outcome is shown in column 3, the lower, risky outcome in column 4, and the probability of the higher outcome in column 5. Column 6 calculates the expected value for the risky choice options, and Column 7 shows the difference between the value of the risk free option and the expected value of choosing the risky option. A participant who chooses the risk free option on row 9, for example, is simply paid $2. However, for a participant who chooses the risky option on row 9 it is played out, for example using dice, and the resulting money (either $2 or $10) is paid out. Since the probability is 0.5, the numbers 1-3 on a six sided die could be assigned to imply the lower payout and the numbers 4-6 the higher payout. The information about the options that is given to participants would usually only include columns 3-5. Adding information from columns 6 and 7 can influence choices in ways that are not desirable, making people act in less risk averse ways than they otherwise would.

A participant who does not care about risk, and who is equally sensitive to all positive monetary outcomes (i.e. has a linear utility function), should pick the option that gives the highest expected value. Thus, if presented with the first row only, such a risk neutral participant would pick the risk free option with a value of $10, over the risky option with an expected value of $6. If instead the participant is given row 6, the choice is between $5 for the risk free option and an expected value of $6 for the risky option, in which case the risky option would be preferred. However, a (sufficiently) risk-averse person would want to stick to the risk free option on row 6, giving up only $1 to avoid the risk of the worst outcome: getting only $2. The more of the expected value of the risky option that a participant is willing to give up, the more risk averse he is. A participant who does not choose the risky option unless the risk free payout is the same or less than the worst payout (task 9 or 10) shows the highest degree of risk aversion. On the other hand, a risk loving participant may want to choose the risky option already in tasks where the risk free option pays more than the expected value of the risky option. It should now be clear that the choice behavior in tasks like this reflects the risk attitude of the participant.

An important consideration is that if a participant is only given one of these rows, we cannot precisely pin down the risk attitude. For example, if we observe a participant making the risk free choice in task 6 alone, all we can say is that the person is not risk loving or risk neutral, but rather risk averse. However, we do not know exactly how averse to risk he is without observing what he would do if offered tasks 7 – 9 as well. Ideally, we would want to observe the participant's choice across all of the tasks and see for which task he abandons the risk free option in favor of the risky one.

There are several different ways in which one may present a series of tasks, like those in Table 1, in order to observe the exact point at which a participant switches from the risk free to the risky task. Each of these alternatives has shortcomings, which we will now discuss.

First, the tasks can be presented exactly as shown in Table 1 (except for removing columns 6 and 7), and the participant asked to make a choice for each task. This method is referred to as a Multiple Price List (MPL). After finishing making all the choices, one of the tasks may be randomly selected to be the only one determining earnings. The task is then played out and the money paid out. This procedure of making payments is referred to as a Random Incentive Lottery Mechanism (RILM) and has the advantage that each task can be given substantial monetary stakes without requiring a very large study budget since only one task is paid for. The problems with this approach are two. First, previous experiments show that many participants are inclined to switch back and forth between choosing the risk free and the risky options. This type of behavior could either reflect a high degree of indifference between the options (they do not care which one they get), perhaps because the difference in payouts is not large enough to motivate a clear choice, or confusion regarding what they are asked to do. Second, even if participants understand the task clearly, if they are the type of people who are either optimistic or pessimistic over probabilities, it is inappropriate on theoretical grounds to use a RILM payment protocol.[5] However, despite these problems this approach is popularly used in the literature.

Alternatively, the tasks can be presented one by one, and immediately after each task the payment is actualized. This is referred to as the Pay All Sequentially (PAS) method. This seems to remove a lot of the switching behavior, but it implies that the participant is not making every choice under identical circumstances, because the wealth of the participant is changing across tasks. Thus, the data analysis needs to incorporate the changing wealth in order for the estimate of the risk attitudes to be unbiased. When presenting tasks in a sequential manner it is, however, important to also vary the order in which tasks are presented to see if there is an influence on choices from the order and avoid any biases in

[5] Cox, Sadiraj, and Schmidt (2012) argue this theoretical point and Harrison and Swarthout (2012) show that behavior is different when using the RLIM protocol.

responses due to such order effects. The drawback of the PAS method is that the payoffs are usually smaller in order for the study to stay within budget, perhaps increasing the degree to which participants express indifference between the choice options.

It is often easier for participants to process the probability and money information in a task if it is visual rather than numeric. Figure 2 shows a popular way of illustrating a choice between two options that are both risky using pie charts. The two colors used in the pie chart represent two possible money payouts. The higher payout is represented in dark blue and the lower payout in light blue. For the left option (A) the higher payout is $100 and the lower payout is $50, as indicated next to the left pie chart and below it. For the right option (B) the higher payout is $1,000 and the lower is $10. The size of the colored areas of the pie chart is a visual representation of the probabilities. In addition, below the pie charts there is also text indicating what die rolls that result in what payoffs. In this case if the die roll is a 1 or a 2 (for a ten sided die) the participant would get the higher amount, so the probability of this outcome is 0.2 and the area that is dark blue is 20 percent. Colors need to be selected keeping in mind that some participants may be color blind. Two other alternatives for visual representation are shown in Figures 3 and 4. Figure 3a is representing probabilities by the length of each box, and the amounts are shown in numbers. It is easy to change this type of image to represent the size of the amounts in the vertical dimension, as the height of the two boxes. This is done in Figure 3b. Figure 4 then illustrates the use of images of money bills to represent the money earnings. In this case the probabilities are 0.5 in each cell, illustrated by the areas of the yellow and blue colors being the same size. There is no systematic study of which visual representation that works best.

Figure 4 also illustrates an alternative type of task for measuring risk attitudes, referred to as the Binswanger method (Binswanger (1980)). In this type of task the participant is shown six different options that vary in both expected value and risk, but with a simple 50/50 probability assignment over outcomes. The 50/50 probability is familiar to most people because it corresponds to a coin toss. The participant is asked to choose one of these six options. While this type of task is easily understood by most participants, in many implementations in the literature it suffers from lack of statistical power to distinguish various degrees of risk aversion. This is due to the particular payoffs that studies have been using and is not an intrinsic problem for this type of task. Thus, if using this mechanism, the researcher needs to pay particular attention to what payoffs are being used, and not just adopt those found in the literature. In the example shown in Figure 4, subjects with a wide range of risk attitudes would all be predicted to pick the same option with $300/$60 payoffs. While such poor choices of payoffs has been a common problem in the literature it can be somewhat corrected by use of different prizes, such as [($100,$100), ($140,$80), ($180,$60), ($220,$40), ($260,$20), ($280,$0)].[6] Further, since the Binswanger method uses only probabilities of 0.5 it is not suitable to estimate optimism and pessimism.

Another method for eliciting risk attitudes is to present participants with a portfolio choice option.[7] In this type of task the participant is again presented with two options: a safer and a riskier one, but instead of selecting one of the two to play out, he can choose how many times out of a total number, say ten, he wants to play the safer vs. the riskier one. Say you offer the participant a choice between a risk free option that pays $5 and a risky option that pays $20 or nothing with an equal chance, i.e. a 0.5 probability. The participant is then given an option to select how many times, out of say 10 total, to be paid according to the risk free or the risky option. For the risky options they must all be actualized with just one die roll so that the outcomes are perfectly correlated. A participant who chooses to be paid 8 times from the risk free option and 2 times from the risky one, and who then rolls the die so that the risky option pays the high prize of $20, would be paid 8 times $5, i.e. $40, plus 2 times $20, for a total of $80. This method has the

[6] Scaling the prizes up or down will not affect the ability to predict risk attitudes, although smaller prizes give respondents less incentive to think carefully about their choices.
[7] Such a method for introduced by Gneezy and Potters (1997).

advantage that it is possible to identify a fairly precise range of risk attitudes with fewer tasks. Instead of the ten that are needed in the first method discussed above (the MPL), it is possible, with the right combination of probabilities and payoffs, to use only three with this method.

All of the task designs shown so far have used stylized, lottery like tasks. It is also possible to introduce contexts so as to generate a more familiar choice setting. However, when doing so one must recognize that the subjects may change the interpretation of the task in ways that are not expected or observable to the researcher. If so, the addition of the context adds a confound to the task that will bias the findings.(reference here to Doing it both ways) A particularly interesting context for route choice is generated when participants make the decisions in simulated environments. We will discuss this in a later section.

<u>To summarize:</u> When eliciting risk attitudes (as curvatures of the utility function) we argue that the method MPL with RLIM suffers from theoretical issues and seems to confuse field participants. PAS appears to have fewer issues with confusion, and does not suffer from the theoretical problems of RLIM, but requires the analyst to properly account for the endogenously changing income. The stakes that can be used in the PAS task are more constrained by the study budget which makes it more likely that extraneous influences affect behavior, or that participants express indifference. The Binswanger method is very easy to understand, but many payoffs used in the literature do not generate the required statistical power. Finally, the portfolios method uses fewer tasks than the MPL to identify the risk attitude, and appears somewhat easier for participants to understand than the MPL (Charness and Viceizsa (2012)). It is advisable to perform a pilot test of various methods before attempting a full-blown elicitation of risk attitudes. Whichever method is used, multiple tasks are necessary in order to properly identify risk attitudes.

ii. Task instructions

The instructions that are given to participants are important both because if they are not properly crafted they can generate confusion, but also because they can lead to a loss of control over the data collection if they do not contain enough detail or if they are not presented in identical ways to each participant. It is rarely possible to achieve both complete control over confounds in the data collection and transparent and easily understood instructions. This is because the participant’s perception of the task can be affected by a great many tangential features of the study session. Often words and phrases are interpreted differently by different participants, and to minimize the range of possible interpretations the instructions need to be detailed and explicit. However, with additional detail the risk of confusion is increasing as the text is more difficult to follow and attention can drop off with longer texts. Again, a pilot testing of various instructions is advisable before going to the full-blown study.

Here is a sample instruction text from a study that uses a PAS payment protocol and pie chart visual displays. The instructions were read to the participants, plus they could read along, but the tasks themselves were performed on a computer.

> “You will be shown two pairs of lotteries on the computer screen. You will see one pair at a time. Here is an image of what the computer screen will look like:
>
> {An image of Figure 2 was inserted here}
>
> On each of the screens you will be asked to make a choice between the two lotteries. You may select either one of them. Each lottery has two possible prizes, a lower one shown in lighter blue and a higher one shown in darker blue. The prizes differ across the pair. In the image above, the left lottery has one prize of $50 and another prize of

> $100. These are just illustrations; the actual prizes will not be this large. The pie chart above the text that shows you the prize displays the chances you have of getting that prize.
>
> The lotteries will be played out using a ten-sided die. The area that is darker blue shows the chances of getting the higher prize, here $100. This will happen if the result of the die roll is either a 1 or a 2. In this case it is 20% of the area, thus a 20% chance, or 2 out of 10. The area that is light blue shows the chances of getting the lower prize of $50. This will happen if the result of the die roll is anything other than a 1 or a 2. It is an 80% chance, or 8 chances out of 10.
>
> You can read the pie chart for the lottery on the right in the same way. The darker blue area is again 20% and corresponds to the chances of getting the higher prize, which is $1000 for this illustrative lottery. The light blue area is again 80% and corresponds to the chances of getting the lower prize, which is $10 in this illustration.
>
> On each screen you will be asked to choose between pairs of lotteries like these. We will play out each lottery as soon as you have finished making your choice. We will then keep track of your cumulative earnings that will be paid to you at the end of the session.
>
> So that you may understand your task better we will first go over a practice round before we start the ones for which you will get paid. In the practice you will not be paid."

Instructions are best given while the participant is actually making practice decisions. Notice that the instruction sample says that it does use payoffs that are very different from those that they will see in the actual tasks. The reason for this procedure is primarily to avoid setting expectations on earnings that will affect behavior, but it also serves to make it clear that the examples are just for practice. The use of a ten-sided die to actualize lotteries is quite common, and most participants do not seem too skeptical about the fairness of this somewhat less familiar die than the common six-sided one.

It is good practice to always allow the participant some practice decisions before engaging in the task that is of interest to the study. Usually there are no money payments for the practice tasks. However, it is a good idea to allow participants to experience one paid task before responding to the ones that will be analyzed, because actually experiencing how the consequences of the choices affect ones pocket book has an effect on behavior.[8]

It is common practice to read the instructions loud to a group of participants while they follow along in their own printed copy. However, when working with non-students it is sometime preferable if the assistance is very personal, so assigning enumerators to each participant can be recommended. Thus, additional enumerators should be in the room to help with rewording or summarizing the instructions for participants who have trouble understanding.

iii. Eliciting optimism and pessimism

Risk attitudes can also derive from optimism or pessimism over probabilities, as discussed earlier. In order to identify how risk attitudes rely on both payoff sensitivity and pessimism it is necessary to add tasks to those that are required to identify the former only. Say you are using an MPL such as the one presented in Table 1. One easy way to modify this is to simply change either the payoffs or the probabilities or both in an added task. However, the changes have to be done in such a way as to make it possible to mathematically identify the additional parameter, following the same logic as solving systems of equations. For many combinations of probabilities and payoffs it can be shown that there is more than

[8] Dixit, Harb, Martinez-Correa and Rutstrom (2013) demonstrate the presence of such earnings effects.

one combination of degrees of income sensitivity and optimism that generates the same choice pattern. If this is the case, all parameters cannot be identified, and when estimated the resulting parameter value could just be nonsense, due to over-fitting the model.

Several of the mechanisms reviewed earlier can be adjusted in this way to allow identification of both components of risk attitudes. The MPL, the sequential presentation, and the Portfolio approach are straightforward and both payoffs and probabilities can be adjusted. The Binswanger method relies on all probabilities being 0.5, thus it is not suitable to fulfill this need.

iv. Eliciting subjective beliefs

As stated in section 3.b route choices and departure time choices depend on both the risk attitudes and the perception of the likelihood of events. In some traffic contexts it may be true that the probabilities over outcomes are known to the drivers, as perhaps they are when the congestion conditions have been stable for a long period and/or intelligent information systems are precise and reliable. However, this is rarely the case, and is a particularly bothersome assumption when predicting responses to changes in traffic systems, such as introductions of congestion pricing. When important changes occur it is reasonable to assume that probabilities are no longer known with precision. It is therefore important to also be able to elicit the beliefs that drivers have over congestion and travel times.

There are two types of tasks that can be used to elicit beliefs: betting tasks and scoring rules. In both cases the participant is presented with an uncertain event that has (monetary) consequences, but for which they are not sure of the probabilities over various outcomes. In order to give them incentives to be as precise as possible the researcher must know the probabilities because the rewards need to be shaped around how close to the true probabilities the expressed beliefs are.

In a betting task the participant is asked to place a bet on possible outcomes. The task could, for example, be designed around a simulation of a traffic network that the participant is familiar with, or perhaps a new traffic network that is being contemplated. Or if the study has access to travel time data on field routes the task could be designed around those. The choice task would then be to bet on whether or not route A would be slower than route B for a given origin-destination. By offering them opportunities to bet at different odds, it becomes possible to infer what they believe the probabilities to be.[9] The question that comes to mind is what money the participant should use to place bets with: his own money or money he receives from the research study? If the participant is asked to use his own money, there is a risk that the inference is confounded by variations in income and wealth that influence the ability and willingness to bet. Further, in the US at least, this would be considered gambling which is a regulated activity. The alternative is to give the participant money with which to gamble. This controls for the willingness and ability to place bets generally, and avoids most regulatory restrictions.

In a scoring rule participants are presented with a set of options of how to get paid, conditional on the true, but to the participant not exactly known, travel time. The simplest scoring rule is referred to as a linear one. Table 2 displays an example of a linear scoring rule. The task is to choose a row to get paid by. If the participant chooses row 3, for instance, he will be paid $8 if the travel time is indeed less than 10 minutes, or $2 if it is 10 minutes or more. Obviously, somebody who chooses this row must be fairly confident that the travel time is less than 10 minutes, but not completely sure. On the other hand, somebody choosing row 9, and who would choose to be paid $2 if the travel time is less than 10 minutes and $8 otherwise, must be fairly, but not completely, confident that the travel time is 10 minutes or more. An alternative to the linear scoring rule is the quadratic one, where they payoffs depend on the square of the probabilities, as shown in Table 3. The difference between them is primarily that if participants are

[9] See Andersen, Fountain, Harrison, Risa Hole, and Rutstrom (2012).

risk neutral, they should choose the first or last row in the linear scoring rule no matter what their beliefs are, while they should choose the row corresponding to their true beliefs in the quadratic one. The weakness of the quadratic rule, on the other hand, is that the foregone earnings by misreporting probabilities close to 0.5 are fairly small so participants will be indifferent between correct and slightly incorrect reports.[10]

One complication in eliciting subjective beliefs is that the task is a risky one: how much money you get is not certain. Thus, the choice in the task depends not only on beliefs but also on the participant's risk attitude. Anytime there is a need to elicit beliefs, the task needs to be complemented by risk elicitation tasks, such as those reviewed in section 3.b.i., and the estimation of the beliefs needs to be performed while controlling for the risk attitudes.

The simplest estimation approach would then estimate the propensity to select a lottery option (in the risk elicitation task) as a non-linear function of the payoffs and probabilities in that task, and do so jointly with estimating the propensity to select an option in the belief elicitation task as a function of payoffs in that task. Because both choices are made by the same individuals, statistical techniques that account for the correlation in error terms between the two estimations should be used. The most general approach is once again SML, where the likelihood function would be defined over both the risk elicitation task and the belief elicitation task.

v. Eliciting subjective value of time

A final important factor in route choice behavior is the subjective value of time, an important determinant to the net value of making a trip. In the field, when driving, the value of the drive depends both on the trip value, i.e. the purpose of the trip, and the cost of the trip, consisting of any road prices paid and the value of the time spent driving. Additionally, there may be costs involved in late or early arrivals. Most of these will vary depending on which route is selected. In our discussions of risk attitudes earlier we used money as an illustration of outcomes that the participant may care about. We also made the point that using money in tasks to elicit risk attitudes and subjective beliefs have the advantage that the researcher knows how the participant views the value of the outcome, removing confounds over unobservable subjective valuations that could lead to biased inferences. However, neither the value of time nor the costs of early or late arrivals are generally observable. We therefore need to device instruments that can be used to infer these values. In this discussion we will focus on the subjective value of time.

A complicating matter for this exercise is the fact that the value of time during a morning commute to work is not the same as the value of time of a participant during a research study session, unless the latter takes place as part of the morning commute. If, for example, we invite participants to a research study involving making route choices in a driving simulator, even if the simulator models the morning commute conditions precisely, it cannot capture the value of time during an actual commute. This is because the opportunity cost of time during a morning commute is not the same as that driving in a simulator in a special research session. The concept opportunity cost of time captures the possible alternative activities that somebody could be engaged in during that time. Thus, for somebody commuting to work it could capture the value of the work that could have been done by arriving to work earlier, or the value of lingering over a cup of coffee reading the news if leaving from home later. In a special research session, however, the participant has set aside a total time for participation which has some opportunity cost in terms of what other activities are foregone. However, once in the session, if doing a driving task with route choices that have simulator travel time consequences, the opportunity cost of that travel time is measured only in activities that could have been undertaken as part of the research session, not those that

[10] An example of instructions for a quadratic scoring rule that can be used to elicit beliefs over a context with more than two outcomes is included in an appendix.

could have been undertaken instead of being in the research session. Because of this, tasks for eliciting subjective value of time must be performed in the natural driving context. The task must be carried out in the field. The most reliable way of recording route choices in the field is to equip the participants' cars with GPS recorders, but an alternative, less expensive, method is to ask them to keep written diaries.

Fortunately, field driving situations offer many variations in travel times that cannot be influenced by the driver and can therefore serve as a natural field trial. Thus, by observing route choices in the field under various travel time conditions, and under conditions that are likely to match the type of opportunity costs that are relevant for the research question, we can infer the subjective value of time. If we know the beliefs that a driver has over travel times, and their attitude to risk including optimism, observing their route choices in the field allows us to infer their value of time. By adding variations in monetary consequences of route choice, we can translate these values into monetary ones.

Small, Winston and Yan (2005) provides a nice example of how a natural field trial can be used to infer value of time. They survey drivers on SR91 after the introduction of a priced lane and ask which lanes they used at various times. They assigned students the task of driving the route at various times, and used those drives to calculate travel time characteristics for both the tolled and the untolled lanes. They then regressed the reported lane choices on the toll, the mean travel time difference, and the difference in the variance of the travel times.[11]Alternatively, one can equip the participants' vehicles with GPS recorders and use the recorded drives both for their observations on route choice, but also for characterizing routes in terms of travel times. GPS recorders are not entirely reliable, however. For example, they are not able to differentiate between choices of lanes that are right next to each other, and recordings sometimes fail due to weather conditions or the presence of obstacles that interfere with the satellite signals. On the other hand, interview approaches suffer from respondents inability to recall their choices accurately and from strategic or social motivations that respondents may have in giving answers that are not true.

vi. Using simulators

One way to introduce a natural context is to program a driving simulator or a traffic simulation and observe how the participants make choices in such environments. Simulators offer a technology with the same rigorous controls over confounds as other experimental tasks but with naturalistic cues and background tasks that make them particularly relevant for policy settings. This experimental design does more than simply present respondents with a contextual frame since the driving task in the simulator requires effort, attention and real time choices, all important aspects of a field situation. One may expect behavioral differences between choices made in simulator environments and the more stylized lottery type tasks described earlier since these are cognitively quite different tasks.[12] It is possible that the lottery task involves slower cognitive modes using explicit deliberation, and that the simulated driving task involves faster cognitive modes using emotions and heuristics. Mukherjee (2010) offers insights into how dual processes, such as deliberative and heuristic processes, may apply to decisions under risk and uncertainty. He suggests that people vary in their disposition to use either of these cognitive processes, and that task construction can directly affect the weight that either gets in the valuation of a prospect. Further, the choice task in the driving simulator may feel more familiar, at least to participants who have little or no familiarity with the stylized lottery task, but who drive a lot. Familiarity may reduce noise in the data, but could also trigger specific heuristics and thus shift decisions in systematic ways.

[11] In fact, Small, Winston and Yan [] estimate not only value of time, but also risk attitudes from the same data. This is possible if there is enough exogenous variations in the tolls and travel time profiles.
[12] Effects like these are proposed by various dual process models in the psychological literature (Chaiken and Tope [1999], Stanovich and West [2000], Kahneman [2003], Evans and Frankish [2009]).

If the goal is to use the driving simulator environment to elicit risk attitudes the driving task needs to be designed with the same controls that are used in the lottery tasks we presented earlier. Therefore, participants must know the value of each possible outcome and the probabilities with which the outcome is actualized. For example, one cannot use a driving task where the outcomes depend on the actual travel time since that will vary across participants due to their varying skill in driving in the simulator. Instead the task can be designed such that the value of the outcome depends entirely on an exogenous random event. Let us present an example of such a design.

Participants are presented with a choice of two routes that will take them from an origin point O to a destination point D in the simulated environment. One route is risk free with guaranteed free flow traffic, but this route has a toll charge. The other route has a risk of congestion with a fixed and known probability. While simulators may not be able to model enough traffic to generate congestion simply due to the traffic volume, it is possible to introduce other congestion causing events, such as slow vehicles or road construction, that increase the travel time. After completing the drive the participant is paid a trip value in dollars. If the participant selects to drive the free flow route the toll is subtracted from the trip value and the participant is paid the net earnings. If the participant selects to drive the route with a probabilistic congestion, a delay cost is subtracted from the trip value, but only if there is congestion. Before starting to drive, a card is drawn from a deck that has ten cards, some of which indicate that congestion occurs and the others that it does not, representing the congestion probability. The participant must know the composition of the card deck, because the congestion probability must be known when eliciting risk attitudes. The researcher loads up the simulator scenario corresponding to either congestion or non-congestion, but makes sure the participant does not know which scenario has been loaded. The participant starts driving and eventually reaches a point where the route choice must be made. There should be nothing in the simulation scenario that indicates whether or not congestion will occur until after the participant has committed to a route. As soon as the participant gets to the point in the simulation where congestion can occur, and notices what the case is, it becomes clear if there is a delay cost or not and the earnings are fully revealed.

The simulator task is very similar to the lottery tasks. Let us take the example shown in Table 1. The trip value would correspond to the high risky outcome of $10, the congestion probability would be 0.5 so half the cards in the deck would indicate congestion and half would indicate no congestion. The delay cost is $8 so in the case of congestion the net earnings on the potentially congested route is $2. The column for the risk free option shows how the earnings decrease as one moves down the rows, corresponding to increasing toll charges. As we discussed earlier, the ideal case would be to have each participant make 10 drives corresponding to each of the rows in order to identify the risk attitude with some precision, although it is important to then consider whether to pay them for one of these 10 drives, selected randomly, or pay them for all drives sequentially.

Apart from eliciting risk attitudes, driving simulators can also be used to study how participants update their beliefs about travel time as they gain experience.[13] In such a design the composition of the card deck should not be known to the driver, but the composition should be kept the same throughout all drives. It may also be useful to allow the driver to draw a number of cards (with replacement) before starting the drives so as to form a prior belief about the likelihood of congestion. As discussed earlier, driving simulators are not useful for eliciting value of time, since the opportunity cost of time in a laboratory setting is very artificial. They are also not useful for eliciting beliefs about travel times in field settings, since it is unlikely that the simulator will be able to match the field driving situation well enough. They are, however, useful for eliciting risk attitudes, optimism/pessimism, and belief updating habits.

[13] The Appendix includes examples of instructions for both types of simulator tasks.

vii. Efficient design

By now it should be clear that it is often important to have many variations in parameter values for the tasks to precisely elicit the characteristics needed. Say for example that you find a need to have ten probabilities and three payoff parameters in the lottery task, and in addition you want to elicit beliefs over travel times for all five commuting days in a week and for each of ten different departure times on those days. If every participant were to respond to all these tasks they would have a total of thirty lottery tasks and fifty belief tasks. If you also wanted to vary the rewards in the belief task in order to test for robustness of the expressed beliefs, you would have even more tasks to implement. Fortunately, you can reduce the dimensionality of the design by recognizing that you may not need to implement every single combination of parameters for all participants. For example, in the lottery task you may assign each probability to only two of the three payoff parameter sets, instead of to all three, as illustrated in Figure 5. Thus, you reduce the number of required conditions to twenty, from the original thirty. However, make sure that you have some variation in each row parameter across each column, and similarly for each column parameter across each row. In addition, you may assign some parameter combinations to a subset of your sample and others to another subset, as long as each subset is large enough to detect or reject systematic effects.

viii. Surveys: Collecting additional characteristics

Risk attitudes, subjective beliefs, and value of time are all expected to vary across individuals. To understand how these variations correlate with demographic characteristics of drivers it is important to complement the ARCT tasks with standard survey instruments. These surveys can include broad characteristics such as gender, age, income and ethnicity, but also driving habits, such as use of toll roads or car pooling habits. The Appendix includes some sample survey questions.

ix. Example of a complete design

In a recent study on behavior under congestion pricing, a design was implemented to elicit risk attitudes, subjective beliefs and value of time. Here we will only briefly review how some of the tasks were designed, but more details can be found in Rutström (2011). The payment protocol used in this study was "pay all sequentially". Participants came to four separate meetings and money payments were executed at the end of each meeting. After each task, the payment consequences were actualized and recorded along with cumulative earnings in a manner that made the earnings very transparent to the participants. Participants were recruited from four different commuter routes: East and West SR408/SR50 in Orlando Florida, North SR400/SR9 and North I75/SR41 in Atlanta Georgia.

Each participant was given four lottery tasks to elicit risk attitudes. These tasks differed in both payoffs and probabilities, and were randomly selected from the set of payoffs and probabilities shown in Table 4. The random selection implies that many different orders of presenting these were used. Because only four tasks were given to each participant we had to pool observations across a larger sample than what would have been necessary if they had each received a larger number of tasks. All lotteries were executed on laptop computers that automatically harvested the data. Each lottery was actualized immediately after the choice was made, and earnings (including cumulative earnings) were clearly shown to the participant. Pie chart visualizations were used in the lotteries.

In addition to the lottery tasks, each participant also completed three driving simulator tasks with known probabilities of congestion, thus also for the purposes of eliciting risk attitudes. The probability was kept fixed across the three tasks, but varied across participants and was 0.3, 0.5, or 0.7. Congestion was simulated by a school bus that could pull up on the risky route, and if that happened the participant had to follow the school bus for 3 blocks and it stopped once in each block. No passing was allowed. The maximum payment, the trip value, was paid when there was no late charge or toll charge, and it varied across participants and took the values $4, $5, and $6. The only parameter that varied across tasks for a participant was the amount of toll charged on the non-congested route, and this was randomly selected from the interval $0.50 up to 50 cents short of the drive value.

In order to elicit the beliefs about travel times, a quadratic scoring rule was used, defined over four outcomes each of which was an interval of travel times.[14] The maximum payoff was either $10 or $5, varied randomly across participants. The probabilities over the four outcomes were constrained to sum to 1. Participants were only given tasks about travel times on their own commuting route. A total of twelve departure times were selected for which their beliefs were elicited, defined over day of the week and specific AM and PM times. To avoid response biases due to influences from being shown specific minimum and maximum travel times, six variations of end points were included. Each subject responded only to two belief tasks, randomly selected from the full set.

Due to the limited number of lottery tasks, simulator tasks and belief tasks given to each of the participants it was not possible to estimate risk attitudes or beliefs precisely for each participant individually, but it was necessary to pool the data. Responses to demographic surveys were then used to estimate how risk attitudes and beliefs varied with observable demographics such as ethnicity, gender, age and income.

Finally, in order to elicit value of time each participant was instructed in a field route choice task on their normal commuter route using their own cars. The private vehicles of the participants were equipped with a GPS recorder which could record 1-2 weeks' worth of data. (Some subjects were given two GPS units to make sure that they did not fill up completely.) The routes consisted of an expressway option and an almost parallel local road with 12-15 traffic signals, approximately half a mile apart. Participants were recruited from areas where it would be safe to assume that they were familiar with these routes, and most participants were daily commuters. We observed the binary route choice on a 5 mile segment of these roads and if participants returned with a recording of a drive on either of these routes during specific morning and evening rush hours, without deviating or stopping for the full 5 miles, they received a monetary drive payment of either $2.50 or $5, assigned randomly across subjects. Each participant drove during three periods of two weeks each, and up to 20 drives in each period were counted as part of the study, thus resulting in payments. During the first period they simply received the drive payment for each valid drive. During the subsequent two periods they got the same drive payment, but in addition they were randomly allocated some combination of charges and subsidies on the two routes. These varied from a charge of $4 to a subsidy of $1.50. For example, one assignment could involve a $2 charge on the expressway, and another assignment could involve a $1 expressway charge and a $1 subsidy on the local road. The use of both charges and subsidies made it possible to test whether reactions to price changes differ across these two domains.

x. A typical protocol

The policy context of each research application will determine how many characteristics that should be elicited. If it can be safely assumed that drivers have adequate information about travel times, then subjective beliefs will not need to be estimated. Further, there may already be reliable information on value of time, thus implying that it may be unnecessary to do a task that requires observations using GPS units in the vehicles of the participants. The only characteristic that then needs to be elicited is that of risk attitudes. We will therefore give an example here of a protocol for a trial involving only elicitation of risk attitudes.

---

[14] Mathiesen and Winkler (1996) introduced this rule over more than two outcomes.

The simplest protocol involves meeting a group of participants and giving them a series of lottery tasks. Subjects are asked to arrive at location early enough to allow check-in procedures before giving task instructions. Check-in procedures include checking that the same person is not participating twice, and giving them a Consent Form, as required by Institutional Review Boards. The main parts of the Consent Form can be explained to them and they should be encouraged to ask questions to clarify the form before signing.

Participants are seated, and after everyone has arrived, the instructions for the task are handed out. When everyone has received a copy, the instructions are read out loud. It may be possible for a single task such as this, to read the instructions to a group of participants at the same time. If the task is executed using computers, the participants may be logged in to access any practice task that is done along with the instructions. Questions can be handled in public or in private, but it is essential that the participants are not revealing to the rest of the group what kind of choices they think should be made, since such remarks may influence the decisions of other participants. As soon as the instructions are done and any additional practice tasks carried out, the actual tasks can commence. Any assistance that is given to the participants should be completely neutral with respect to what choices they should make. To the extent possible, the text from the written instructions should be used for clarifications and only minimal language changes be allowed.

Either before or after the lottery task, participants can be given any other survey instruments that are included. If given before, they should be collected before reading the instructions for the lottery task so as not to generate a distraction for the participants, and to ensure that all participants finish the survey.

After the lottery tasks are finished, record sheets, if used, should be checked and collected. A receipt with the total earnings should be filled in and signed by the participant so the researcher can be reimbursed from the funding source. Payments should always be made in complete privacy so as not to reveal to other participants what kind of decisions were made. All record sheets, or computer records, should use an anonymous subject id and not include any personal information to safeguard participant anonymity.

If many tasks are used in the study they can be given out sequentially, building in breaks for the participants. Breaks are especially important if using driving simulators since participants may become nauseous if staying in the simulator for too long. As a general guide, a task in the simulator should not last longer than five minutes without a break. It may even be necessary to break the study up over several meetings. If so, it is necessary to store contact information in secure ways, and to keep track records of how to merge the responses for the same individual across multiple tasks and meetings.

c. Implementation

i. Sampling

Sampling and recruiting for an ARCT follows the same standards as any other field study. The sample should be representative of the target population, and to the extent possible, avoid sample selection biases. Thus, announcing the purpose of the study, such as the purpose of evaluating congestion pricing policies, may cause a stronger incentive to participate among some groups than others, and if these groups differ in, say, risk attitudes, the results will be biased. It is also important to have a fairly substantial fixed money payment for participation to avoid the more risk averse members of the population from seeing the study as too risky to join, and also to make sure to encourage participation from participants with a high cost of participating. Harrison, Lau and Rutström (2009) show evidence that sample selection based on risk attitudes occurs, but that a higher participation compensation can minimize that. Special efforts to reach minority and low income populations may be necessary, depending on their trust in the institution conducting the study.

ii. Enumerators

Enumerators must be solidly trained in how to conduct controlled, randomized trials. Participants must be guaranteed anonymity and poorly trained enumerators can easily violate this condition by speaking too loudly or by using somebody's name. When assisting participants in understanding task instructions or questions, poorly trained enumerators can introduce behavioral confounds in many ways. For example, it may be tempting to give examples of types of choices that the enumerator believes an average person might make, but this can induce behavior in that direction. While sometimes it is necessary for the enumerator to use simplified language to explain a task, it is important that all enumerators are trained to do this in the same way so that all participants get similar explanations. Otherwise the range of interpretations of what the task is can be very wide, and most importantly, unobservable to the researcher.

The training of enumerators can involve the following steps. First, give an overview of the research session with a quick introduction to the surveys and tasks used. Second, go over specific rules of conduct, such as those discussed in the previous paragraph. Give some examples of actions and language that should not be used. Three, make sure all enumerators study and rehearse the instructions and scripts. Four, in a series of follow-up meetings enumerators should practice conducting all steps of the study using each other, and other individuals, as the participants. These practice events should include feed-back and discussions of ways to approach the tasks and instructions. To the greatest extent possible, all enumerators should participate together so that they converge on a similar behavior. A minimum of two such practice events covering all surveys and tasks should be conducted, but ideally there should be four to five to achieve convergence and stability in enumerator behavior. No matter how much training is given, close supervision is required while enumerators get used to working with actual participants. However, such supervision cannot be intrusive or interfere with the participant's sense of privacy or anonymity.

Apart from the specific training in conducting the sessions enumerators will also have to pass Institutional Review Board training and tests to be fully certified.

iii. Creating a lab in the field

Working with field participants may require researchers to move outside of their usual institutional settings, such as university laboratories. Finding locations that are convenient for participants is important, but equally important is to have facilities that allow you to store some of your equipment securely, especially if you are using simulators that may be difficult and time consuming to transport. If the study involves participants who live or work in many different places it may be necessary to have more than one location. If so, it is important to not perform all sessions in one location and then move to another, but to either run sessions at multiple locations at the same time, or to vary locations within each week. If conducting the sessions serially across multiple locations the risk is that the field situation of participants may change and that such changes confound any analysis of differences in behavior across locations. One problem with conducting sessions at multiple locations at the same time is that oversight of the enumerators necessary to make sure they behave in the same way at all locations becomes difficult. It also requires a much larger staff of enumerators.

iv. Preparations and logistics

Before going out in the field to conduct the sessions it is necessary to prepare all the materials such as surveys, instructions and softwares. It is easy to underestimate the amount of time that is needed for these tasks, especially if the study involves several dimensions of individual characteristics so that many tasks are involved. It may even be necessary to break up the tasks across multiple sessions, consequently adding logistical effort for keeping track of each participant's data and information without violating privacy and anonymity.

4. Summary and Conclusions

Augmented Randomized Controlled Trials offer an approach to evaluating proposed congestion pricing schemes that is less costly than full-scale field trials. Compared to the current practice of stated preference studies, it gives the respondents incentives to respond in ways that better match the behavior they would display in field route choices where they do face significant consequences. Compared to the current practice of revealed preferences, ARCT can directly manipulate the availability of routes and the prices charged, thereby allowing for a wider set of contexts in which to observe route choices than what is possible under the natural circumstances of revealed preferences. ARCT therefore brings together the ability of the stated preference approach to allow a wider range of circumstances to be studies, with the ability of the revealed preference approach to implement real consequences.

In this guidebook we have discussed how to elicit the human factors that enter route and departure time choices in traffic situations with congestion and congestion pricing. We have identified the main factors as consisting of risk attitudes, perceptions of congestion probabilities, and value of time. Without understanding these factors, predictions of how drivers will respond to new congestion pricing initiatives can be suffer from serious biases. Based on how these factors are theorized to matter for decisions, we showed ways to design tasks to give participants in the study which will allow the researcher to infer these factors. We discussed the importance of designing tasks that properly identify all the factors, and that allow a great deal of control over other influences on behavior that could confound the inferences. Further, we discussed a number of practical issues in implementing the design to collect data using field participants, such as commuters.

Table 1: Example of Task for Measuring Risk Attitudes

Task Number	Risk Free Option	High, Risky Outcome	Low, Risky Outcome	Probability of High Outcome	Expected Value of Risky Option	Risk Free Minus Expected Value of Risky
1	$10	$10	$2	.5	$6	$4
2	$9	$10	$2	.5	$6	$3
3	$8	$10	$2	.5	$6	$2
4	$7	$10	$2	.5	$6	$1
5	$6	$10	$2	.5	$6	$0
6	$5	$10	$2	.5	$6	-$1
7	$4	$10	$2	.5	$6	-$2
8	$3	$10	$2	.5	$6	-$3
9	$2	$10	$2	.5	$6	-$4
10	$1	$10	$2	.5	$6	-$5

Only columns 1-5 are shown to respondents, and they are asked to make a choice between the risk free and the risky option for each task.

Table 2: A Linear Scoring Rule

Row	Probability	Payment if Travel Time is Less Than 10 Minutes $10*probability	Payment if Travel Time is 10 Minutes or More $10*(1-probability)
1	1	$10	$0
2	.9	$9	$1
3	.8	$8	$2
4	.7	$7	$3
5	.6	$6	$4
6	.5	$5	$5
7	.4	$4	$6
8	.3	$3	$7
9	.2	$2	$8
10	.1	$1	$9
11	0	$0	$10

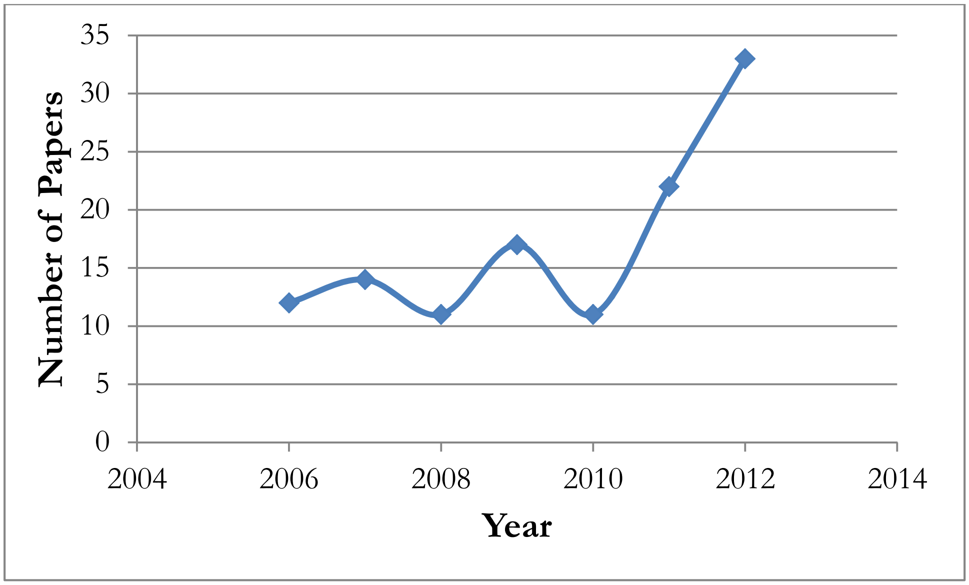
Table 3: A Quadratic Scoring Rule

Row	Probability	Payment if Travel Time is Less Than 10 Minutes $10*probability^2	Payment if Travel Time is 10 Minutes or More $10*(1-probability)^2
1	1	$10	$0
2	.9	$8.1	$.01
3	.8	$6.4	$.4
4	.7	$4.9	$.9
5	.6	$3.6	$1.6
6	.5	$2.5	$2.5
7	.4	$1.6	$3.6
8	.3	$0.9	$4.9
9	.2	$0.4	$6.4
10	.1	$.01	$8.1
11	0	$0	$10

Table 4: Parameter values for sample lottery task

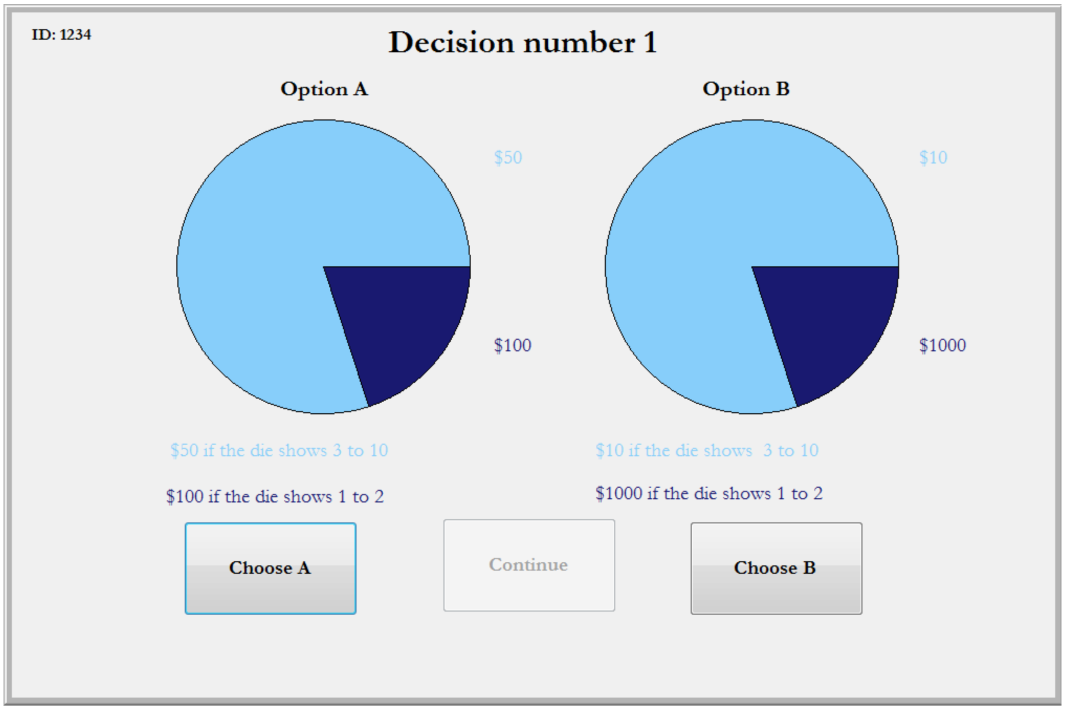
<table>
<tr><td></td><td colspan="6">Safe lottery</td><td colspan="6">Risky lottery</td></tr>
<tr><td></td><td colspan="3">Low Payoff</td><td colspan="3">High Payoff</td><td colspan="3">Low Payoff</td><td colspan="3">High Payoff</td></tr>
<tr><td>Payoffs</td><td colspan="3">$2</td><td colspan="3">$3</td><td colspan="3">$0.25</td><td colspan="3">$4</td></tr>
<tr><td>Payoffs</td><td colspan="3">$2</td><td colspan="3">$3</td><td colspan="3">$0.25</td><td colspan="3">$5</td></tr>
<tr><td>Payoffs</td><td colspan="3">$2</td><td colspan="3">$3</td><td colspan="3">$0.25</td><td colspan="3">$6</td></tr>
<tr><td>Payoffs</td><td colspan="3">$4</td><td colspan="3">$6</td><td colspan="3">$0.50</td><td colspan="3">$10</td></tr>
<tr><td>Probabilities of High Payoff</td><td>.1</td><td>.2</td><td colspan="2">.3</td><td>.4</td><td colspan="2">.5</td><td>.6</td><td colspan="2">.7</td><td>.8</td><td>.9</td></tr>
</table>

Figure 1: Experiments on Congestion Pricing



Result of Google Scholar trends of papers in: “Traffic Congestion” and “Experimental Economics”

Figure 2: Visual Display of Lottery Choices



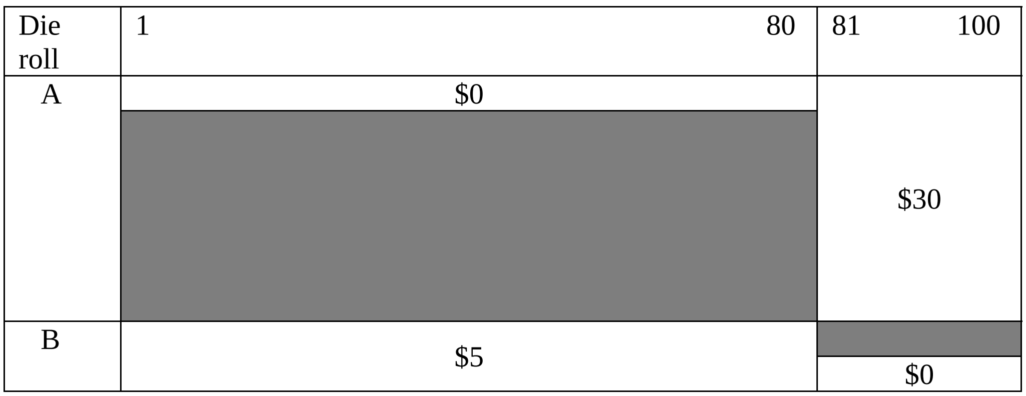
A series of pairwise choices like these can be shown to respondents, one at a time. For each one they choose either A, the left one, or B, the right one.

Figure 3a Visual Display of Probabilities and Prizes in a Lottery

Percent chance	1 80	81 100
A	$0	$30
B	$5	$0

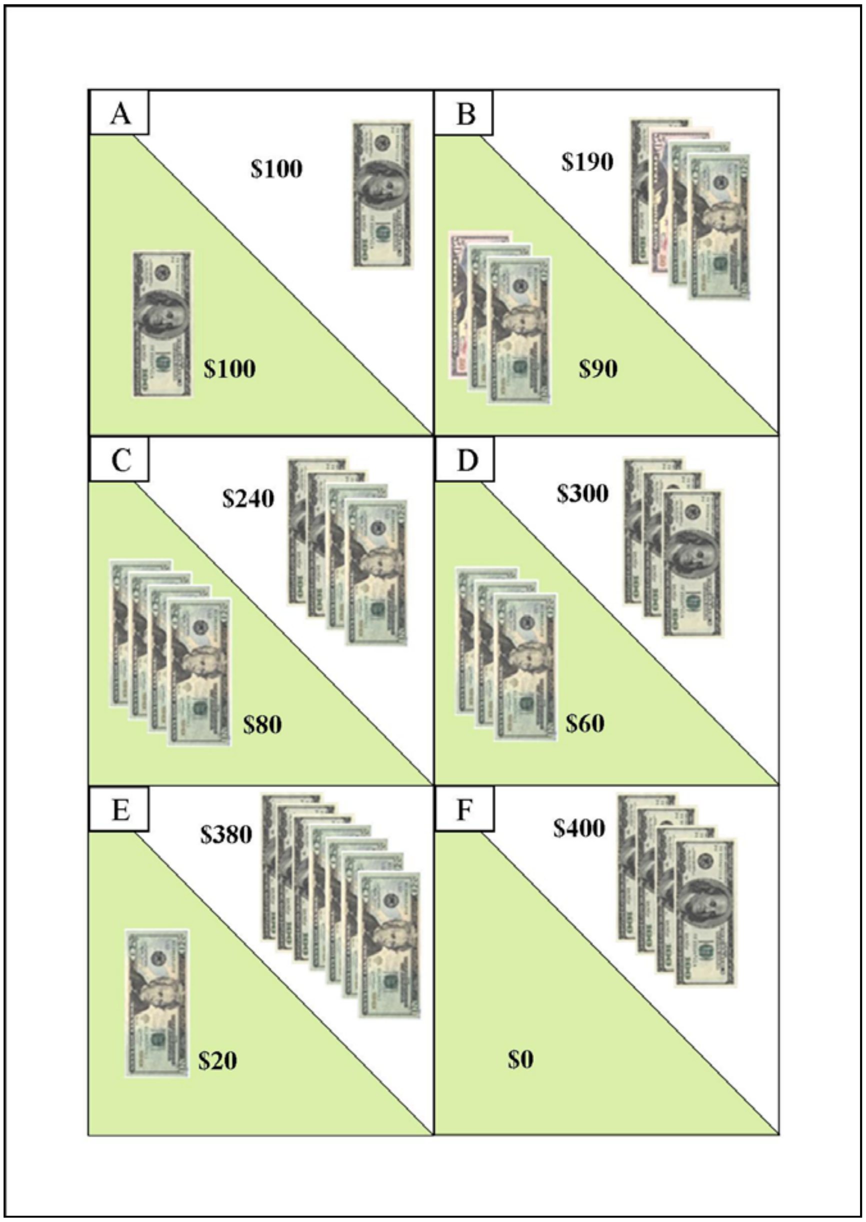
Probabilities are illustrated through the length of each box, and the monetary amounts are shown inside each box. Option A is a 20% chance for $30 and an 80% chance of nothing, while option B is an 80% chance of $5 and a 20% chance of nothing. This design is due to Beattie and Loomes (1997)

Figure 3b Alternative Visual Display of Probabilities and Prizes in a Lottery



The height of each box illustrate the magnitude of the prizes. Probabilities are illustrated as in Figure 2a.

Figure 4: Display of Lottery Choices with Money images



This type of display was originally used in Barr (2003).

Figure 5A blocking design of a lottery task

probabilities	Payoff set 1	Payoff set 2	Payoff set 3
.1	XXXXXXX		XXXXXXX
.2		XXXXXXX	XXXXXXX
.3	XXXXXXX	XXXXXXX	
.4	XXXXXXX	XXXXXXX	
.5	XXXXXXX		XXXXXXX
.6		XXXXXXX	XXXXXXX
.7		XXXXXXX	XXXXXXX
.8	XXXXXXX	XXXXXXX	
.9	XXXXXXX		XXXXXXX
1.0		XXXXXXX	XXXXXXX

Cells marked with “XXXXXXX” are implemented and empty cells are skipped.

References

Andersen, Steffen, John Fountain, Glenn W. Harrison, Arne Risa Hole, and E. Elisabet Rutström, "Inferring beliefs as subjectively imprecise probabilities", *Theory and Decision*, 73, 1, July 2012, 161-184.

Andersen, Steffen, Glenn W. Harrison, Morten I. Lau, and E. Elisabet Rutström, "Eliciting risk and time preferences," *Econometrica,* (76) 3, May 2008, 583-618.

Anderson L., Holt, C.A. and D. Reiley (2008). Congestion Pricing and Welfare: An entry experiment. In Cherry, Koll, Shogren (eds), Environmental Economics, Experimental Methods, chapter 15, 280-92.

Barr, Abigail (2003), "Risk pooling, commitment, and information: An experimental test of two fundamental assumptions", Working Paper 2003-05, Centre for the Study of African Economies, Department of Economics, University of Oxford.

Battalio, R.C., Kagel, J.H., Jiranyakul, K., "Testing between alternative models of choice under uncertainty: Some initial results*."* *Journal of Risk and Uncertainty* 3, 1990, 25–50.

Beattie, J. and Graham Loomes, "The impact of incentives upon risky choice experiments", Journal of Risk and Uncertainty, 14, 149-162.

Binswanger, Hans P., "Attitudes towards risk: Experimental measurement in rural India,"*American Journal of Agricultural Economics,* 62(3), August 1980, 395-407.

Charness, Gary and Angelino Viceisza, "Comprehension and risk elicitation in the field: Evidence from rural Senegal," *IFPRI discussion papers 1135*, International Food Policy Research Institute.

Chmura, Thorsten, and Thomas Pitz, "An Extended Reinforcement Algorithm for Estimation of Human Behaviour in Congestion Games," Bonn Econ Discussion Paper 24/2004, Bonn Graduate School of Economics, December 2004a.

Chmura, Thorsten, Thomas Pitz and M. Schreckenberg "Minority Game – Experiments and Simulations of Traffic Scenarios," Bonn Econ Discussion Paper 23/2004, Bonn Graduate School of Economics, December 2004b.

Cox, James C., Vjollca Sadiraj, and Ulrich Schmidt, 2012, "Paradoxes and Mechanisms for Choice under Risk", *ExCen Working Paper 2012-06*, Experimental Economics Center, Georgia State University.

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

Daniel, T. E.; E. J. Gisches and A. Rapoport (2009). Departure Times in Y-Shaped Traffic Networks with Multiple Bottlenecks. The American Economic Review Vol. 99, No. 5 (Dec., 2009), pp. 2149-2176

Denant-Boemont L. and Fortat V. (2013). Snowball Effect and Traffic Equilibrium in a Market Entry Game: A Laboratory Experiment. Transportation Research Board Annual Meeting 2013, Washington D. C.

Dixit, Vinayak V., Rami C. Harb, Jimmy Martinez and Elisabet E. Rutström (2013), "Measuring Risk Aversion to Guide Transportation Policy: Naturalistic Tasks and Respondents", *DBEL-WP1301*, Dean's Behavioral Economics Laboratory, Georgia State Univerisity, December 2013.

FHWA, "Congestion Pricing, a Primer," *Office of Transportation Management*, Federal Highway Administration, Washington. October 2008.

Fleischfresser, C. "Top 10 most congested cities in U.S.", http://www.smartplanet.com/blog/transportation/top-10-most-congested-cities-in-us/1026, 2011.

Gneezy, U., and J. Potters "An Experiment on Risk Taking and Evaluation Periods," *Quarterly Journal of Economics*, CXII, May 1997, 631-645.

Harrison, Glenn W., "Hypothetical Bias Over Uncertain Outcomes" in *Using Experimental Methods in Environmental and Resource Economics,* John A. List, ed., Northampton, MA: Elgar, 2006.

Harrison, GlennW., Erik Johnson, Melayne M. McInnes, and Elisabet E. Rutström, 2005. Risk Aversion and Incentive Effects: Comment. *The American Economic Review* 95, 897–901.

Harrison, Glenn W., Morten I. Lau, and E. Elisabet Rutström, "Risk Attitudes, Randomization to Treatments, and Self-Selection into Experiments," *Journal of Economic Behavior and Organization* 70, 2009, 498-507.

Harrison, Glenn W., and E. Elisabet Rutström, "Risk Aversion in the Laboratory," in J.C. Cox and G. W. Harrison (eds.*), Risk Aversion in Experiments* (Bingley, UK: Emerald, Research in Experimental Economics, Volume 12, 2008).

Harrison, Glenn W., and Todd J. Swarthout, "The Independence Axiom and the Bipolar Behaviorist", *ExCen WP 2012-01*, Experimental Economics Center, Georgia State University, 2012.

Holt, C.A., Laury, S.K., 2002. Risk Aversion and Incentive Effects. *The American Economic Review* 92, 1644–1655.

Kahneman, Daniel and Amos Tversky, 1979, "Prospect Theory: An Analysis of Decision under Risk", *Econometrica*, (47(2), March, 263-291.

Li, Zheng., and David Hensher, "Congestion charging and car use: A review of stated preference and opinion studies and market monitoring evidence", *Transport Policy*, issue 20, 2012, 47-61

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

Nielsen, O.A. and G. Jovicic (2003). The AKTA Road Pricing Experiment in Copenhagen, paper presented at the 10th International Conference on Travel Behaviour Research, Lucerne, August 2003.

Quiggin, John, “A theory of anticipated utility”, *Journal of Economic Behavior and Organization* 3(4), 1982, 323–43.

Quiggin, J. *Generalized Expected Utility Theory. The Rank-Dependent Model*. Boston: Kluwer Academic Publishers, 1993.

Rutström, Elisabet E., “Experiments on Driving Under Uncertain Congestion Conditions and the Effects on Traffic Networks from Congestion Initiatives”, Report to FHWA EAR program, *DBEL_WP1101*, Dean’s Behavioral Economics Laboratory, Georgia State University, November 2011.

Schrank, D. and Lomax, T., “2012 Annual Urban Mobility Report,” *Texas Transportation Institute*, College Station, TX. December 2012.

Selten, Reinhardt, Thorsten Chmura, Thomas Pitz, S. Kube, and M Schreckenberg (2007), “Commuters route choice behavior”, Games and Economic Behavior, 58, 394-406.

Small, Kenneth A., Clifford Winston, and Jia Yan, “Uncovering the Distribution of Motorists Preferences for Travel Time and Reliability”, *Econometrica*, 73(4), July 2005, 1367-1382.

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

Tversky, Amos and Daniel Kahneman, “Advances in prospect theory: Cumulative representation of uncertain prospects”, Journal of Risk and Uncertainty, 5(4), October 1992, 297-323.

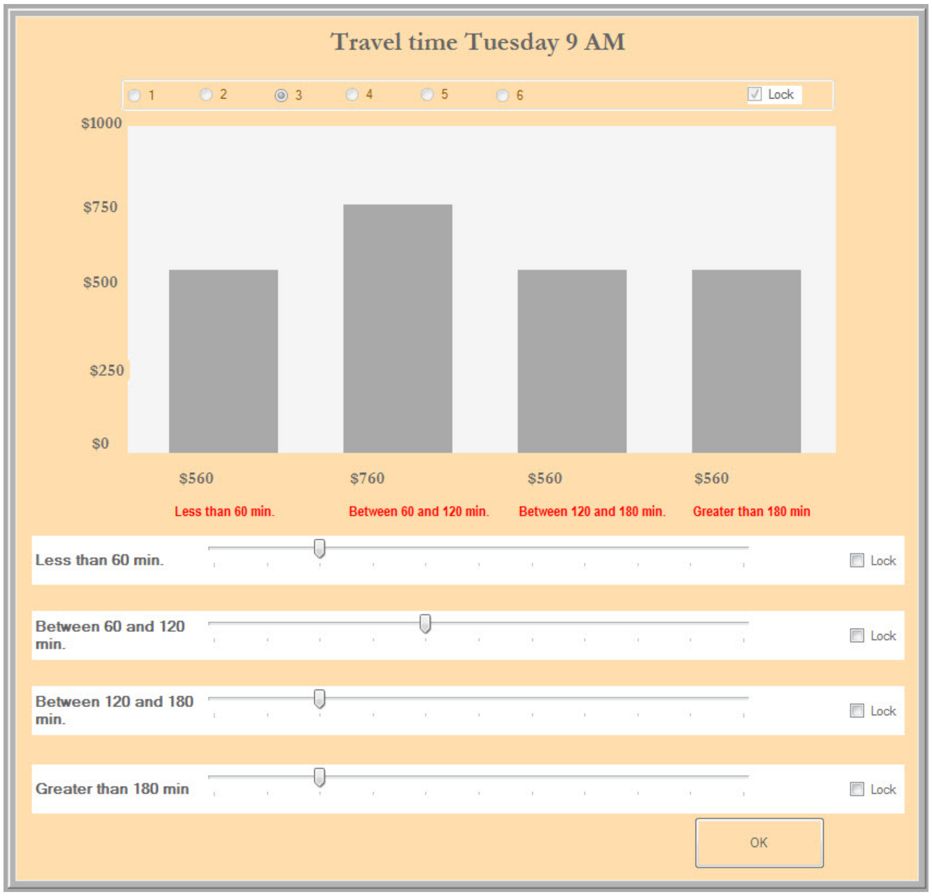
Ziegelmeyer , A, F. Koessler, K. B. My and L. Denant-Boemont (2008). Road Traffic Congestion and Public Information. Journal of Transport Economics and Policy, Volume 42, Part 1, January 2008, 43–82.

APPENDIX: Instructions for various tasks

Instructions for Belief Elicitation of Context with More Than Two Outcomes:

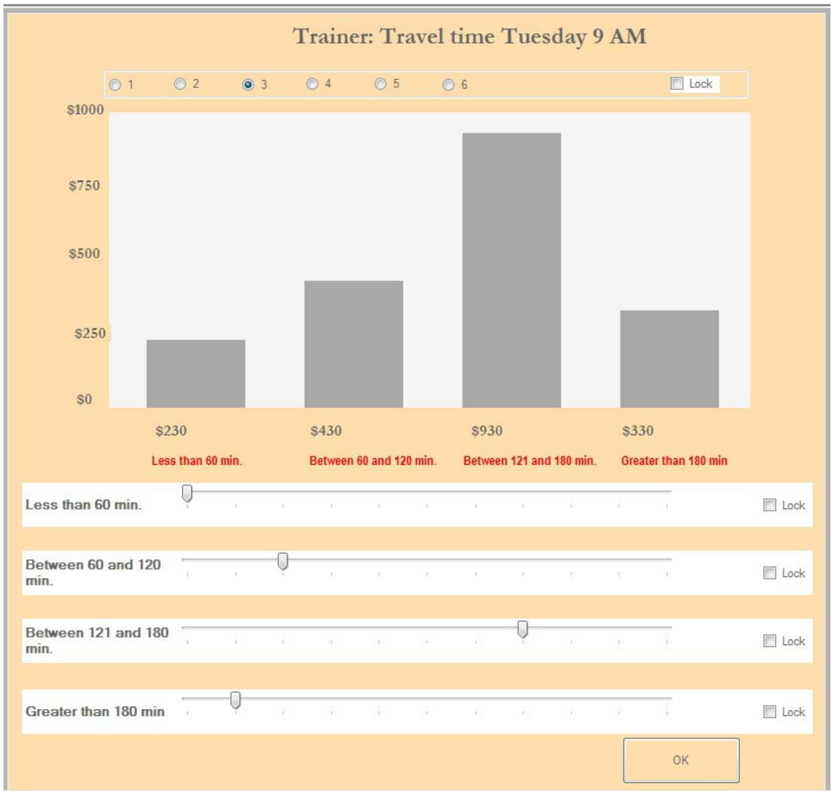
In this task you will be paid according to how long it takes to travel on SR9 (Roswell Road) and SR400 from I-285 to Northridge Road. Your earnings depend on the DIFFERENCE in travel time between these two routes. As all of our participants return after driving on these two routes during the next two weeks we will download the data. Then, we will calculate the travel time for each driver. We collect all southbound drives that start at Northridge Road between 7 and 7:30 am, 8 and 8:30 am, and 9 and 9:30 am on Mondays and Tuesdays. We also collect northbound drives that start at I-285 between 4 and 4:30 pm, 5:30 and 6 pm, and 7:30 and 8 pm on Tuesdays and Fridays. For each of these 12 sets of travel times we then calculate the average travel time across all participants. Then, we take the difference between the average travel time on SR9 and the average travel time on SR400. You will be asked to make decisions that impact your earnings for two of these time periods, randomly selected.

You will be shown a page on the computer which looks like this. On the top of the page you can see which of the 12 time periods is selected. In this illustration it is Tuesday at 9 am. You may get a different one. There are also six buttons numbered 1 to 6 that will be explained later.



The top part of this computer screen displays earnings that you may make. In this illustration they are shown as much higher than in the actual tasks just to make it clear that it is an illustration. Earnings depend on what the actual travel time is. The height of each bar measures the dollar amounts that are shown on the left. The bottom part shows you four sliders that you can move around to change how much you get paid. In this illustration the top slider determines how much you get paid if the actual travel time difference is less than 60 minutes. The left bar in the top part of the image shows you that this is $560. The next slider is moved to determine how much you get paid if the actual travel time difference is between 60 and 120 minutes. The earnings are displayed in the second bar from the left at the top and are $760. The next two sliders similarly are used to choose how much you want to get paid if the travel time difference is between 120 and 180 minutes, or greater than 180 minutes respectively.

As you move the sliders the bars showing your earnings will move too. The image below shows an example where the first slider has been moved to the left. Thus, lowering the earnings you get if the travel time difference is less than 60 minutes from $560 to $230. The third bar has been moved much further to the right. The travel time difference is between 120 and 180 minutes and you would be paid $930.



If you want to continue to adjust only two or three of the sliders, you can lock the fourth one in place by checking the “Lock” box on the right.

The six buttons above the bars select the number of minutes that will be shown to the left of the sliders. Before making your decisions you will roll a regular 6-sided die to select one of these 6 options. After clicking the appropriate button you will see the number of minutes that your earnings are defined over.

You can now practice using this computer program before making any decisions for money.
In the practice you will not be paid. After practicing you will be given two of these tasks for money.

Instructions for Simulator Driving Task that Elicits Risk Attitudes

Respondents are first shown a video of the simulator route choice task, before they are given these instructions about the task.

Practice Drives

Please familiarize yourself with the controls of the simulator. The steering wheel and pedals act just like the comparable features of a normal automatic vehicle although they feel a bit different.

Do you have any questions?

<u>In this first practice drive</u> you are going to take 7th Ave to work. This is for practice only and you will earn no money based on this drive.

<u>In this second practice drive</u> you are going to take 9th Ave to work. Do you have any questions?

How do you feel? Do you have any feeling of nausea at all?

Would you like to get up and move around, perhaps have some water?

Remember from the video that sometimes when you take 9th Ave there will be a school bus pulling up from C street. This bus stops at every intersection from C to F Street. If this happens, you and the other cars in front of you will have to also stop.

<u>In this third practice drive</u> you will once again take 9th Ave and we will make sure that there is a school bus present so you can see what happens then.

Paid Drives

Your first paid simulator task will be to drive to work from home, using either 7th or 9th Ave. You will do this three times. Which way you take is your choice. I will not tell you which way to go. While you are waiting at the red light at the 7th Ave intersection you can make your choice.

By driving to work, you will be able to earn a $_________ daily wage. If you end up behind the bus on 9th Ave., you will be late for work and will lose $_________ of your daily wage, thus being paid only $.25 cents. If you do not end up behind the bus, you will arrive to work on time and receive your full $________ wage. If you take 7th Ave. you will receive your $_______ wage minus a toll that is charged for this route. Thus there is no risk of getting to work late if you take 7th Ave.

Notice that no matter how fast you drive you will always be late for work if you get stuck behind a school bus.

On 9th Ave. there is a chance that a school bus will be on the route with frequent stops. You will not know for sure, however, if there is going to be a bus.

Here I have three identical decks with ten cards. _______ of them have a bus on them, and _______ of them do not. Please take a look at the cards to make sure that this is the case.

Before you drive you will select a card from one of the decks without seeing whether it has a bus on it or not. I will then look at it and start the simulation. If the card shows me a bus I will start a simulation that has a school bus, and if the card does not show a bus I will start a simulation without a bus. You will be able to view the card after you have finished all three of your drives, but not before. We will use a new deck for each drive so that the proportion of cards with a bus stays the same.

On 7th Ave. there is no bus, but a toll must be paid if you take this route.

Here are three other decks of cards, labelled "Tolls." As you can see in the first deck here, we have tolls ranging from $_____ up to $_____, the second deck has tolls ranging from $_____ up to $_____, and the third deck has tolls ranging from $_____ up to $_____.

I will separately shuffle all three decks now, one deck at a time and then I will place them in front of you face down. So you will see three shuffled decks.

Please select <u>three toll cards,</u> taking one from each deck. Turn them over. This is the amount of money you will have to pay out of your wages if you take 7th Ave. You will be doing this drive three times, each with a different toll.

We will now use a 6-sided die to select the order in which you will drive with these tolls. If the first roll shows 1-2 you will have the low toll first, if it is 3-4, the medium toll and 5-6 the high toll. If the second roll is 1-3 your second toll will be lower of the two remaining.

We will record the order of these tolls for use in a later task.

I will now shuffle these 10 Bus cards. You will choose one, but I will not reveal the card to you until after you have driven. Thus, when you drive, you do not know in advance if there is a bus on 9th Ave..

This handout shows you a summary of the task. Please review it before we start.
Do you have any questions?

Are you ready for your first drive? Do you know what the toll is?

Please take a Bus card.

Are you ready for your second drive? Do you know what the toll is? Please take a Bus card.

It is time for the third drive. Do you know the toll? Please take a Bus card.

Handout

- Pick a Bus card – do not turn it over
- Start at home
- Your daily wage is $____
- If you take 7th Ave. pay the toll on the toll card – it is subtracted from your wage
- If you take 9th Ave. and there IS a bus your wage is reduced to $____
- If you take 9th Ave. and there is NO bus you get your full wages
- There is never a bus on 7th Ave.
- You may choose either 7th or 9th Ave
- You may make a different choice each time you drive

Do this 3 times, each time with a different toll.

Instructions for Driving Simulator Task Where Congestion Risk is Unknown

It is time again to drive in the simulator. It will be similar to the previous drives you did for money, but your earnings will be calculated differently.

For these paid drives, you will NOT know in advance whether there is a bus on 9th Ave or not. We have a deck of 100 cards here that contain a mix of cards with and without a bus. You will not know how many of these cards have a bus on them until you have finished all your drives today. The proportion of cards that have a bus is not at all the same as last time. There could be many more cards with a bus, or much fewer ones. After you are done with all the drives you may look at all the cards.

You will be allowed to draw 10 cards from the deck before we start so that you can get a feel for how many cards that may have a bus on them.

On 7th Ave. you will pay a toll that you select from our deck of Toll cards. This toll will stay the same for all your drives. Let us select the toll card now. You may review the cards before we shuffle them.

Your earnings in this part depend on you arriving to work on time. Of course, this depends on how long it takes to do the drive. Your wage today is $______. Any toll or charges for travel time will be deducted from this wage.

- If you take _______minutes and ______ seconds or less to get to work, you will arrive on time and will receive your full $_______ wage.

- If you take longer than _ ___minutes and ______ seconds, you will be late, and your wages will be reduced by $ ________.

The maximum charge you will incur will not exceed your wage of $________, so you will never incur any losses.

Please review the bullet handout for an overview. Do you have any questions?

I will now shuffle the bus cards.
You will now draw 10 cards (one at a time) from the deck with bus cards and look at them. You will return each card to the deck after viewing it. I will reshuffle the deck between each draw.

We record the outcome of your card draws on the record sheet.

Handout

Draw a bus card without looking at it.

- Start at home
- Drive to work
- If you drive 7th Ave. you pay a toll
- If you drive 9th Ave. there is no toll
- Earnings:
    - Wage $_______
    - Deduct $_______ if drive takes longer than ___ minutes and ___ seconds
    - Pay a toll if you take 7th Ave.

Save the bus card for later verification and recording.

Do this five times. Each time you draw a bus card from a complete, but identical, deck of cards. There is a separate deck for each time you drive, but all of these decks of cards are identical – they have exactly the same proportion of bus cards. You will be allowed to verify this, if you like, after completing all of your drives.

Demographic Questionnaire

In this survey most of the questions asked are descriptive. We will not be grading your answers. Please think carefully about each question and give your best answers
.

1. What is your gender?
   - Male
   - Female

2. What is your Age?
   - 18-21
   - 22-25
   - 26-30
   - 31-40
   - 41-55
   - 56-75
   - Over 75

3. Which of the following categories best describe you? Choose all that apply.
   - White not Hispanic or Latino
   - Hispanic or Latino
   - Black or African American
   - American Indian and Alaska Native
   - Asian Indian
   - Chinese
   - Other Asian
   - Native Hawaiian and other Pacific Islander
   - Some other race
   - Two or more races

4. How would you best describe your household?
[A household is an economic unit. It is defined as a group of persons who live in the same residence and each person contributes to general expenditures. Your household includes your spouse, children or parents who live with you, and all siblings, partners or room-mates with whom you share finances]

- Single under 30 years
- Single 30 – 59 years
- Single older than 59 years
- 2 adults, oldest person is under 30 years
- 2 adults, oldest person is 30 – 59 years
- 2 adults, oldest person is older than 59 years
- Single with children, oldest child 0 – 9 years
- Single with children, oldest child 10 – 17 years
- 2 adults with children, oldest child 0 – 9 years
- 2 adults with children, oldest child 10 – 17 years
- Household with at least 3 adults

5. How many people are there in your household (including your spouse, children or parents who live with you, and all siblings, partners or room-mates with whom you share finances)?

- 1 person
- 2 persons
- 3 persons
- 4 persons
- 5 or more persons

6. What was the total pre-tax income earned in 2010 by all members of your household (including your spouse, children or parents who live with you, and all siblings, partners or room-mates with whom you share finances)?
[Please consider all forms of income, including salaries, income from unincorporated business enterprises, pension scheme contributions, interest earnings and dividends, retirement benefits, student grants, scholarship support, social security, unemployment benefits, parental support, alimony, child support, and other types of income.]

- $15,000 or under
- $15,001 - $25,000
- $25,001 - $35,000
- $35,001 - $50,000
- $50,001 - $65,000
- $65,001 - $80,000
- $80,001 - $100,000
- $100,001 - $200,000
- Over $200,000

7. How many cars are there in your household (including those owned by your spouse, children or parents who live with you, and all siblings, partners or room-mates with whom you share finances)?

- 0
- 1
- 2
- 3
- 4
- 5+

8. How many cars in your household have a manual transmission (stick shift)?

- 0
- 1
- 2
- 3
- 4
- 5+

9. Are you employed (including self-employed)? Select the option which best applies to you.

- Full-time employed
- Full-time student without employment
- Part-time student without employment
- Part-time employed
- Unemployed

9b. (If yes to employed) Does your job involve driving during working hours? For instance delivering pizzas or other goods or cargo, driving to clients for onsite repairs like plumbing, electrical, etc...

- Yes, 3 or more days per week
- Yes, less than 3 but at least 1 day per week
- No

10. How many <u>other</u> adults in your household are employed at least part-time?

- 0
- 1
- 2
- 3
- 4+

11. What has been your primary occupation during the last 12 months?
[Primary occupation is defined as the type of occupation where you spend most of your working time.]

- Farmer
- Other self-employed
- Spouse assisting in family business
- Caring for spouse for medical reasons
- White collar worker
- Professional
- Skilled worker
- Unskilled worker
- Apprentice
- Student
- Retired
- Unemployed
- Stay at home spouse and/or mother
- Other: ___________________

12. What type of residence do you live in?

- Owner-occupied house
- Owner-occupied apartment
- Owner-occupied mobile home
- Rented house
- Rented apartment
- Rented mobile home
- Multi-ownership of residence, cooperative
- Rented room
- Other: __________________

13. What is the zip code of your residence? ___________

14. What is the square footage of your residence? ___________

15. What is your highest level of formal education?

- Less than high school
- GED or High School Equivalency
- High school
- Vocational or trade school
- Two year college degree (AA degree)
- College or university
- Graduate degree

16. What was the highest level of formal education that your father (or male guardian) completed?

- Less than high school
- GED or High School Equivalency
- High school
- Vocational or trade school
- Two year college degree (AA degree)
- College or university
- Graduate degree

17. What was the highest level of education that your mother (or female guardian) completed?

- Less than high school
- GED or High School Equivalency
- High school
- Vocational or trade school
- Two year college degree (AA degree)
- College or university
- Graduate degree

18. Do you currently smoke cigarettes?

- No
- Yes

18A. If yes, how much do you smoke in one day? ________ cigarettes

19. How often do you participate in extreme sports?

[Extreme sports include bungee-jumping, para-gliding, parachute jumping, gliding, rafting, diving and other dangerous sports.]

- Never
- A few times
- Occasionally
- Often
- Every chance I get

20. Have you ever played video games in which you can move around in a 3-dimensional world?

- Yes
- No
- I am not sure

21. Have you ever participated in online virtual worlds, such as Second Life?

- Yes
- No
- I am not sure

22. Have you ever played online video games, such as World of Warcraft?

- Yes
- No
- I am not sure

If “Yes”, name your favorite video games here:

23. How often do you play video or computer games or participate in online worlds?

- Never
- A few times
- Occasionally
- Often
- Every chance I get
- I am not sure

24. Do you regularly play video games that encourage the fast driving of vehicles in a virtual world, for example *GTA*?

- Yes I play such games frequently
- Yes but only on occasion
- No I never or infrequently play such games