Experimental Economics and Choice in Transportation: Incentives and Context

Vinayak V. Dixit, Andreas Ortmann, E. Elisabet Rutström and Satish V. Ukkusuri

January, 2017

This paper reviews the preconditions for successful applications of Experimental Economics methods to research on transportation problems, as new transportation and research technologies emerge. We argue that the application of properly designed incentives, the hallmark of Experimental Economics, provides a high degree of experimental control, leading to internal validity and incentive compatibility. Both of these are essential for ensuring that findings generalize to contexts outside the immediate application. New technologies, such as virtual reality simulators, can generate external validity for the experiments by providing realistic contexts. GPS and other tracking technologies, as well as smart phones, smart cards and connected vehicle technologies can allow detailed observations on actions and real-time interactions with drivers in field experiments. Proper application of these new technologies in research requires an understanding of how to maintain a high level of internal validity and incentive compatibility as external validity is increased. In this review of past applications of Experimental Economics to transportation we focus on their success in achieving external and internal validity.

Affiliations

Dixit is an Associate Professor at the School of Civil and Environmental Engineering at UNSW Australia and Director of TRACSlab.

Ortmann is Professor of Economics at the Australian School of Business at UNSW Australia.

Rutström is a Research Director at the Center for Economic Analysis, Georgia State University, an Honorary Professor at the University of Cape Town, and a Visiting Professor at Örebro University.

Ukkusuri is a Professor in the School of Civil and Environmental Engineering at Purdue University.

This material is based upon work supported by the Federal Highway Administration under Agreement No. DTFH61-09-H-00012.

1. INTRODUCTION

One fundamental goal of transportation science is to understand behaviour of stakeholders relevant to transportation problems. Planners and researchers in transportation science have relied heavily on concepts of traffic equilibrium and relatively untested assumptions about user behaviour. Much of the behavioural data used to test these assumptions has relied on surveys, interviews and focus groups, where the questions pertain either to historic actions or intended actions under various hypothetical traffic circumstances with no consequences (see Hensher 2010 for a comprehensive review). While direct observations of behaviour in the field would be preferred, the often prohibiting cost of undertaking large-scale field studies makes them rare. Experimental Economics (hereafter EE) offers an alternative method that allows direct observations with actual consequences but at lower cost and with the advantage of being able to control and manipulate traffic circumstances and contexts along the lines of stated choice.

Emerging technologies in the area of transportation may transform how research is conducted, and what questions that can be answered. Two important foundations of good research design is internal and external validity, both of which matter to the successful incorporation of new technologies in research. Internal validity refers to the extent to which the causal relationships can be inferred, while external validity is the extent to which these causal relationships can be generalized to other settings (other people, situations, etc.). EE uses incentives that are compatible with Induced Value Theory (Smith 1976, 1982) to implement internal and external validity (through the existence of actual consequences of actions taken). This theory states that to gain experimental control over motivations that influence choices, it is often critical to design reward structures to induce values (or motivations). The reward structures should be designed such that the rewards are the dominant source of motivation during the experiments. The reliance on Induced Value Theory distinguishes EE from survey data collection methods including stated choice methods. While it may be possible to design surveys that induce intrinsic motivations that satisfy these criteria, we restrict this review to experiments that rely on extrinsic, and thus directly observable, motivations.[1] External validity can be further strengthened through the use of natural contexts and participants from populations with relevant experiences. Additional field relevant attributes can be introduced via text, images or visual simulations.

The predominant mode of behavioural data generation in transportation relies on stated choice surveys. The stated choice approach presents participants with a set of choice situations that vary in travel related attributes such as mode of transportation, travel routes, departure and arrival times, as well as travel times, possibly in addition to policy variables such as road pricing. The stated choice approach provides researchers with responses to a large set of circumstances at a low cost. The weakness of the approach is the lack of actual consequences,

[1] For readers that are interested in understanding the difficulties in generating intrinsic motivations that fulfil the criteria we refer to Hensher (2010), Fifer et al. (2014), Cummings et al. (1995), Blumenschein et al. (2008), Champ et al. (2009, 1997), Johannesson (1999), Li and Mattsson (1995), Moore et al. (2010), and Morrison and Brown (2009).

which is likely to lead to both noise and biases in responses. [2] Hensher (2010) and Fifer et al. (2014) discuss these difficulties and propose some solutions. Revealed preference elicitation is an alternative method that relies on actual consequences. Applications of this method frequently rely on natural experiments, i.e. changes in transportation that are not controlled by the researcher (see for example Small et al. 2005). Data collection under revealed preference can use either interviews about past choices or direct collections using GPS technology. However, they cannot be relied on to study multiple policy changes in a comparative fashion, and the cost of collecting such revealed preference data can be high. While stated choice studies may suffer from bias due to the lack of real incentives, revealed preference studies may suffer from poorly defined non-chosen alternatives, lack of significant variance of key variables, and other measurement errors (Hensher 2010). EE is similar to stated choice studies in that circumstances and contexts can be manipulated directly and similar to revealed preference studies in that actual consequences are incorporated. It is worth noting that EE has reached a level of maturity equal to other fields in economics and that it is widely accepted as a tool in economists' methodological tool box[3].

Our review is timely given the advent of innovative communication technologies such as smart phones (Ben-Elia and Ettema, 2011a and 2011b), smart cards (Pluntke et al. 2011) and connected vehicles (albeit currently through OBD ports for instance Hultkrantz and Lindberg, 2011). These technologies create opportunities to conduct controlled, large scale incentivized field studies at a lower cost. They allow researchers to make more reliable observations on actions and conditions, and they make it possible to have real-time communications with study participants. Further, the improvement in Virtual Reality technology has made it possible to conduct immersive and realistic laboratory experiments that have a higher degree of external validity. Furthermore, many concerns about the applicability of EE to transportation has been raised out of misconceptions about the method, and there is therefore a need for clarifications. Our article is meant to fill this gap and to provide an improved understanding both the difficulties in and the value of generating internal and external validity by reviewing lessons from past applications of EE to transportation.

In this review we restrict our attention to studies that rely on Induced Value Theory and that have a direct link to transportation questions. We focus on experimental design factors rather than the appropriateness of the theory or the econometric estimation techniques.[4]

---

[2] Such hypothetical response biases have been demonstrated in Cummings et al. (1995), Holt & Laury (2002), and Harrison & Rutström (2008)

[3] EE has grown exponentially and now includes several dedicated journals, the Economic Science Association with over 500 members, and four winners of the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel (Reinhard Selten, Vernon L. Smith, Elinor Ostrom, Alvin E. Roth). Ortmann (2016) recounts key episodes from the history of Experimental Economics.

[4] We have selected articles for this review based on scopus and google scholar searches with key words "Transportation" and "Experimental Economics". We also emailed the members of the Economic Science Association (ESA) to identify additional papers. We do not claim that we have been able to exhaustively review all existing work in transportation that rely on Induced Value Theory.

2. BACKGROUND AND EVIDENCE

Individual choice experiments in transportation concern *route choice, mode choice, departure time choice, and location choice.* The policy questions that these experiments want to inform include *congestion pricing, public transit funding, and constructing additional lane miles and routes.* There are two primary perspectives in the literature: a focus on the choice of an individual when the attributes of the choice options, including the congestion conditions, are *exogenous*, and a focus on the choices of groups of individuals when they affect attributes such as the congestion level such that these are *endogenous*. Applications with exogenous attributes have been conducted in both the lab and the field, while applications with endogenous attributes have mostly relied on lab environments.

In a lab environment it is possible to generate a high degree of internal validity, i.e. of designing conditions that match those assumed by the theory being tested or by the context assumed. In a field environment this is significantly harder, but it may instead be possible to generate a greater degree of external validity, i.e. conditions that are more natural or familiar to the participants and where the theory may have little to say about how choices are affected by these conditions. We expect that the introduction of new technologies can help bridge the lab and the field environments by bringing more natural conditions into the lab and by allowing a greater degree of control and internal validity in the field.

We will start this section by reviewing experiments in the exogenous environment category and then discuss the endogenous environment.

2.1 Choice in Exogenous Traffic Contexts

The interest in studying individual drivers in exogenous transportation environments is to better understand the decision factors and personal characteristics behind the choices made. With an improved understanding of the sources of heterogeneity in individual decision making it will be possible to test the validity of assumptions used in transportation planning models and to predict how emerging technologies will affect transportation behaviour. Experiments can be designed to generate data for reduced-form decision models, such as Mean-Variance Utility, or for other latent non-linear decision structures. Experiments in this area include lab experiments with stylized tasks, lab experiments with driving simulators or simulated traffic environments, and field experiments.

2.1.1 Field Experiments

Incentives to Influence Behaviour. Significant work has been undertaken to evaluate how incentives can be utilized to influence behaviour in route, departure time and mode choice.

<u>Route Choice and Pricing</u>**.** An example of a field experiment involving route choice is Nielsen and Jovicic (2003). Their study, with 400 volunteer participants, was part of the AKTA[5] project

[5] AKTA is the Danish abbreviation for Alternative Driving and Congestion Charging and was part of the European Union's PROGRESS project (www.progress-project.org).

in Copenhagen, Denmark. Their participants were given an initial monetary budget and were charged varying toll rates reflecting those that were planned in the policy process. A pricing scheme was implemented using virtual cordon rings and pricing zones. After an initial baseline 8-10 week period with no pricing, various pricing schemes were implemented for another period of 8-10 weeks. All trips made during the study were logged using GPS recorders. The participants were paid according to the difference in their driving behaviour between the baseline and the priced periods. Participants were recruited such that their usual driving would take place within the price-manipulated zones. A methodological conclusion from the study, which was designed to have high external validity, is that relying on GPS technology can lead to unanticipated additional effort and costs. Behaviourally, the researchers reported considerable inflexibility on behalf of the participants in changing driving habits, a lack of awareness of alternatives, and considerable heterogeneity in the cost variables that they focused on. The experimental manipulations were very complex, and the main take from this study is the great degree of heterogeneity in what motivates each trip and each participant.

<u>Departure Time Choice and Rewards.</u> Merugu et al. (2009) reports a field experiment in departure time choice. The authors studied commuters who worked at Infosys Technologies in Bangalore. In the experiment commuters were offered incentives to shift to earlier arrival times. Each participating commuter received credits for swiping in at work either before 8 a.m. (1.5 credits) or between 8 and 8:30 a.m. (1 credit). At the end of each week participants with at least 3 credits participated in a draw to select winners who received cash. The mechanism was designed such that more credits induced higher probability of winning, and cash prizes. Once the winners were determined for the week, credits for both winners and non-winners were deducted so as to drop them down to a level immediately below the one they had qualified for. This ensured that the accrued credits over a week did not diminish the impact of incentives for the next week. Over the six months of the experiment the number of bus commuters who arrived before 9 a.m. increased by almost 30%. This field experiment provides a high degree of external validity.

Pluntke et al. (2013) is a field experiment conducted in Singapore with the aim of shifting public transit users from on-peak travel to off-peak travel. The incentives used are based on earning credits for each trip taken during weekdays for both on-peak and off-peak travel, but the off-peak trips earn 3 times the credits of on-peak trips. The credits are redeemable in cash. Members of the scheme can also earn credits by signing up friends and family, and there is a personalized bonus paid out to members who are underperforming in either switching to off-peak travel or in inviting new members. The authors report a 10% shift from on-peak to off-peak travel. They find that those who became members through friends' invitations, and those who chose the raffle reward over the fixed reward, shift more than others. The INSINC program exemplifies a number of factors where caution in design needs to be applied in order to avoid participant strategic behaviours leading to increased program costs. For example, the additional rewards for off peak travel is likely to attract more off peak commuters than on peak commuters to participate in the program, particularly since the details of how credits are given can easily be transmitted with the social recruitment mechanism (participants get bonus credits if they recruit friends and families to participate). In addition, the personal incentives for those who underperform is an example of an incentive scheme that, if not implemented cautiously, can also lead to strategically withholding shifts to off peak hours in anticipation of such underperformance incentives. As with the other field experiments, the familiarity of the transportation setting leads to high external validity. In addition to the common weak internal validity in large field studies, the potential strategic concerns may have eroded internal validity further.

Mode Choice, Information and Rewards. Tørnblad et al. (2014) implemented a field experiment with 327 participants recruited from six companies in Oslo. The authors wanted to test whether information about public transit options, alone or in combination with a small set of free bus tickets, would shift travel choice from cars to buses. While the experimental manipulations were real, the choice observations were self-reported through survey responses, rather than collected using technology. Revealed preference surveys were conducted four weeks prior to introducing the experimental manipulations, five weeks after, and five months after. The study found no significant effect on mode choice for any of the treatments. Again, the study achieved external validity due to the familiarity that participants had with the transportation system prior to participating.

Departure Time, Mode Choice and Rewards**.** A seminal field experiment with high external validity is Ben-Elia and Ettema (2011a and 2011b). It involved 341 Dutch commuters on the Zoetermeer – The Hague corridor over a 13-week period. The purpose was to see if participants would respond to the reward structures offered and either change to different travel times or modes. The experimental manipulation offered the commuters monetary compensation for avoiding the corridor during peak hours. The reward varied between €3 and €7 per trip. The first two weeks involved no monetary rewards, but established a reference travel pattern that determined how the rewards were paid out. This design feature deals with some of the concerns about strategic behavior expressed above. Drivers who used the corridor three times per week during the rush-hours of the reference period could only receive a reward for the third, fourth and fifth day in the week that they avoided the rush-hour. Drivers who used the corridor during the rush-hours of the reference period five times per week were eligible for a reward for all five days in the week. An alternative reward of a Yeti Smartphone was included for those participants who so chose. Participants were given the smartphone up front, and could keep the phone only if avoiding the corridor during peak hours as determined by the reference period drive patterns. Verification of route avoidance was done using in-vehicle installed transponders, electronic vehicle identification, and road-side cameras. Complementary data was collected in the form of daily web-based logbooks and a pre-test survey of travel routines and demographic characteristics. The study found significant changes in travel behaviour.

The strength of these four studies lies in them being conducted on the routes and in the modes that the proposed policy was going to be potentially applied to, thus external validity was high. Since routes were familiar to the participants, any contextual preferences or perceptions that influenced behavior in the study would be the same ones that would influence their behaviour in the field, outside of the study. While familiar field settings generate high external validity it is not necessarily the case that findings are transferable to other traffic regions and driving populations, unless additional information is collected under conditions of high internal validity that will allow estimation of appropriate decision models. In addition, field studies that offer incentives to change behaviour need to be concerned about possible leakage of information about these incentives since they can affect baseline behaviour and result in biased conclusions. Further, such information leakage can result in sample selection whereby a larger portion of individuals who stand to gain from the incentives sign up to participate than the proportion that would result from a random selection. As new technology makes it possible to conduct field experiments at lower cost, it will improve the opportunities of achieving higher internal validity through additional individual level observations, but it may also increase the risk of information leakage and the possibility of strategic bias.

Safety. Speed is an important determinant of accidents since it affects both the number and the incidence of accidents. To handle the externality in speed choice, most societies use a combination of regulations (speed limits), enforcements (fines) and insurance schemes with deductions and bonus-malus provisions.

PAYS: In 2002, Hultkrantz and Lindberg (2011) conducted a field experiment of a Pay As You Speed (PAYS) insurance scheme involving 95 drivers in Sweden. The participants for this study were recruited from a larger vehicle-fleet trial with 1000 drivers involving Intelligent Speed Adaptation (ISA). 250 of these 1000 drivers agreed to have their vehicle fitted with on-board computers that had digital maps, GPS positioning and mobile communication facilities. The technical system informed the driver about the speed limit on a display in the vehicle while an acoustic signal and a flashing light alerted him or her if the speed limit was exceeded. The experiment varied both a monthly endowment of money given to participants from which the penalties would be deducted and the amount of the penalty. The design put a cap on how much penalty could be charged so that each participant was guaranteed to get earnings of at least 75 SEK per month, in essence allowing participants to go bankrupt instead of incurring additional penalties. In principle, such a design can make participants risk loving as they approach the bankruptcy threshold since they have nothing left to lose. The authors report, however, that no one was even close to the bankruptcy threshold. Baseline observations on speeding were obtained from the larger, earlier study, thus making it unlikely that these observations were affected by strategic concerns. The authors were also able to investigate sample selection effects since they were given access to data on speed collected from the remaining drivers with in-vehicle technology from the original ISA trial. There is also a reported “Hawthorne” or “placebo” effect, as speeding was reduced initially, but this reduction was only sustained in the conditions with penalties. In this study, higher penalties do not generate stronger effects, which is in line with experimental evidence that the effectiveness of incentives is often not a linear function (e.g., Holt and Laury 2002, Rydval and Ortmann 2004).

Similar studies were later conducted in Denmark (Agerholm et al. 2008), but with the purpose of testing if financial incentives, such as a discount on their insurance premium, would incentivize younger drivers to adopt an ISA technology. The study offered 30% discounts on insurance premium for drivers aged 18-28 years. Each participant first drove during a 1.5 month long baseline period. During this time there were no signals to the driver indicating speeding from the ISA unit, nor were there any other interventions aimed at reducing their speed. Following the baseline period each driver was randomly assigned into one of four conditions, varying incentives and information, all employing ISA. This “ISA period” lasted 4.5 months. If the car exceeded the speed limit by more than 5 km/h, the driver received a verbal warning. The third warning resulted in penalty points. Each penalty point deducted 7 Euro cents from the 30% discount. There was also a cap on the cumulative penalty so that the insurance premium was capped at the non-participation level. Their results were based on 38 participants: The combination of information and incentive resulted in the highest reduction in speeding, followed by information provision and the incentive-only treatment. This indicates that incentives without information provision might have minimal impact on speeding behaviour. The control group, which received no penalties or verbal warnings, increased their speeding during the “ISA Period” compared to the baseline, consistent with an initial Hawthorne effect that wears off eventually.

Bolderdijk et al. (2011), in collaboration with five Dutch insurance companies, carried out an incentivized PAYS experiment with 141 young drivers in the Netherlands. The participants could receive up to 50 Euros discount monthly: up to 30 Euros for not exceeding the speed limit, 15 Euros for reduced mileage, and 5 Euros for avoiding driving during weekend night time hours. Excessive speeding (driving over 20% of the speed limit) and speeding during weekend night-time hours resulted in additional penalty points. The study was divided into four phases. The first and last phase had no incentives so only observations on speeding and mileage were done. There were two treatment conditions: a gain frame and a loss frame, in addition to a control group with no incentives. The study reports that the PAYS scheme significantly reduced speeding on all types of roads. However, there was no significant impact on mileage driven, or reduction in night time driving. Furthermore, there was no significant difference in behaviour between presenting the incentives as a loss or gain.

These three studies show that one need to worry about possible “Hawthorne” effects, thus allowing a sufficiently long study period for these to disappear before drawing inferences. Simply signalling speeding to the driver has no effect, but adding penalties is successful. Incentives are not significantly different across gain and loss frames. Caps on penalties were instituted to guarantee participants positive net earnings, but such design features could result in risk loving behaviour.

Strategic concerns with field experiments. An early field experiment by Bohm that was never published but which is discussed in Bohm (1984, p. 136) offers an important related lesson. Bohm undertook an experiment on a new bus service to a major hospital in Stockholm. The purpose was to find out how valuable this new bus service would be by eliciting the willingness to pay from participants who would be affected by the added route. The experiment was to provide the bus route for a 6-months trial period and during that time riders would pay for the service. As it turned out, when the experiment started nobody turned up to use the bus service due to a union boycott. The union was concerned that in this city with publically funded transit, evidence of private payments would lead to political decisions to keep the new bus line privately funded. The important lesson here is that any experiment, field or lab, is linked to activities and concerns that potential participants have outside of the experiment, and these influence both the decision to participate, leading to sample selection, and decisions made in the experiment, which potentially confound the results. Many field experimental protocols require the participants to be informed about the incentive structure and the mechanism, and therefore experimenters need to carefully consider any strategic concerns that might arise. For example, individuals may behave differently if they believe that their behaviour could shape policy.

These field experiments have a number of things in common. They all had a control condition and treatment conditions; they all used clear monetary incentives in the treatment conditions; they were conducted in participants' natural environments; and most of them used technology to collect the observations directly. While field experiments can have a high external validity, they provide less control of other factors in various ways. For example, the number of experimental manipulations is limited as is the ability to observe many behavioural factors that determine choice. Lab experiments, in contrast, may have weaker external validity but allow a greater degree of control of other factors. It is possible to use lab experiments to perform preliminary hypotheses tests in a more controlled and less expensive environment prior to conducting a field experiment. At other times lab experiments can be complementary to field experiments when conducted on the same participants, providing measures of agent characteristics that are difficult to identify in the field.

Complementary Field and Lab Experiments. Andersen et al. (2014) conducted a route pricing field experiment in the Orlando and Atlanta areas with daily commuters, complemented with lab experiments involving synthetic choice tasks as well as simulator based choice tasks to measure decision characteristics of the participants. 497 commuters participated in four lab experiment sessions and three field driving periods over an 8-10 week period. For the field driving experiment route choice as observed via GPS recorders on a 5 mile segment of the participants' usual daily commute between home and work, both in the morning and the evening. The restriction in route length and times of the day was implemented to restrict the variations in trip purpose and trip value, that otherwise would confound the inferences. The routes are important commuter routes, with one being an expressway and the other being an arterial road with intersections and traffic signals. Monetary incentives were used for both the driving task and the belief elicitation task. In the former task drivers were paid either $2.50 or $5 for each drive (with a limit of 10 drives per week), and were then charged various prices for accessing either of two routes. Adding lab experiments with tasks to elicit beliefs about travel times allowed the estimation of value of time and value of travel time reliability based on participants' perceptions rather than on actual travel times. Further lab experiments with lottery choice tasks were used to identify utility over money, such that a multi-attribute utility function could be estimated. The additional lab tasks lead to significant adjustments in the estimates of value of time and value of reliability, indicating that relying on actual travel times and on linear relations between time and money may lead to biased inferences.

We summarize the findings from field experiments that have studied choices in exogenous contexts in Table 1.

Table 1: Summary of findings from field experiments with exogenous context

Study Questions	Key Findings
	High external validity, low internal validity. GPS technology improves reliability of data collection. ISA technology used for enforcement of speeding penalties.
Incentives: Route Choice	Lack of awareness of alternatives Heterogeneity in motivations Inflexibility to change habits Unanticipated costs and effort with GPS technology
Incentives: Departure Time and Mode Choice	Mixed findings on effects on departure time choice No significant impact of incentives was found on mode choice. Concerns with strategic behaviour
Safety	Information alone has no effect on speeding but adding penalties does. Gain and loss frames are not significantly different Initial Hawthorne effects require longer study periods Successful use of ISA technology
Strategic concerns	Strategic concerns can generate biased responses and can influence sample selection
Measurement using complementary lab experiments	Travel time perceptions do not match actual travel times. More flexible utility functions lead to different inferences.

2.1.2 Laboratory Experiments with Driving Simulator

Route Choice and Pricing. Dixit et al. (2015) used route choice tasks in driving simulators to estimate risk attitudes and compared these to risk attitudes estimated from more standard stylized lottery choices. The paper is a methodological contribution comparing risk attitudes not only between these two types of tasks, but also across a sample of participants recruited from daily commuters and a sample of university students. The stylized lottery choices were presented as pie charts with two colours, each representing one of two possible outcomes: a high monetary prize and a low monetary prize. The proportion of the pie that was covered by each colour represents the probability, similar to a lottery wheel. In addition, text under each pie chart explained which numbers on a ten-sided die corresponded to which prize, the high and the low. Such a ten-sided die was used to play out the lottery. Participants saw two of these pie charts at the same time, with the one on the left being less risky than the one on the right. The task was to choose which one of these two to play out for money. Prizes varied from $0.25 to $10. The choice task in the simulator was designed to be very similar to the stylized lottery, to make them as comparable as possible. The risk facing participants was that of congestion on one of two available routes in the simulation. The perception of risk was controlled by drawing cards from decks with known proportions indicating congestion. The value of time was induced with money. While there were initial differences in risk attitudes, both across type of tasks and type of participants, very minimal experience was required to make them the same. The conclusion is that tasks that are easier to implement, such as stylized lotteries, and participants that are less costly to recruit, such as university students, can be used to learn a lot about risk attitudes in more involved, contextual tasks for individuals who are familiar with that context.

Tsang (2015) reports on a lab experiment with participants recruited from the commuting population in Atlanta and Orlando. The experiment offers the participants a route choice task conducted in a simulated driving environment where the choice is between a safe route that has a toll charge and a route that may get congested, but where the probability of congestion is not known. Participants get an initial random sample of ten observations from which they can form their prior beliefs about the congestion risk. After these prior observations each participant makes ten drives in a simulated environment. These drives are hypothesized to influence the beliefs drivers hold over the congestion risk, but the experience is asymmetric in that the experience is only informative about the congestion risk if the driver chooses the risky, congested route. If the driver chooses the safe route, he cannot observe whether the non-selected route is congested or not. The author finds that participants in the low congestion risk condition use the risky route more than those in the high congestion risk condition and are therefore able to update their beliefs through experience. Of course, in many field traffic settings GPS or traffic radio ameliorate these asymmetric information conditions, thus weakening the external validity (e,g,, Devetag and Ortmann 2007; see also Selten et al., 2007, and Chmura and Pitz, 2007, which we discuss below).

Safety. There are numerous studies that have used driving simulators to study safety; there are, however, only two studies whose design has been based on Induced Value Theory.

Subjective Beliefs of Crash. Dixit et al. (2014) conducted an incentivized laboratory experiment with student participants to investigate whether risk attitudes or beliefs about the likelihood of crashing affected crash propensity. A driving simulator was used to model a left-turn task with an oncoming queue of vehicles. All values were induced: the value of making a successful turn was set at $10, the cost of a crash at $5, and the cost of delaying the turn by one other vehicle in the oncoming cue was set at $1. If a driver successfully turned in the very first gap of the oncoming queue the earnings was $10, with every extra gap he waited reducing earnings by a dollar, and crashes reducing earnings by $5. The benefit of waiting was that the gaps between the cars in the oncoming queue got larger, thus lowering the risk of crashing. Before making the first turn, the drivers were given tasks to measure their accuracy in the perceptions of gap sizes and to get familiarity in operating the driving simulator. After the first turn (which was paid), drivers were given the opportunity to make three more practice turns without earnings, before making a final, paid turn again. Finally, participants were paid on successfully turning four times in randomly selected gaps between on-coming vehicles, this task was used to measure their ability to drive in the simulator. The key observation collected in the main turning task was which gap they selected to turn in. In addition they were given simple lottery tasks of the kind employed in Dixit and Denant-Boemont (2014), to measure their risk attitudes. Controlling for their risk attitude, the gap they chose to turn in is interpreted as reflecting their beliefs about the relationship between gap sizes and crash propensities. The authors report that the inferred beliefs depend on both the level of drivers' experience (as measured by comparing the first to the second paid turning task) and skill (as measured in the final turning task with randomly selected gaps). The variation in the propensity to crash was better explained by variations in the beliefs about crashing than risk attitudes.

Insurance demand. Accident forgiveness is often considered as a type of "premium insurance", it protects the insured against a premium increase in the case of an at-fault accident, and provides rewards in terms of coverage and premiums for those who remain accident-free. Liu (2012) conducted laboratory experiments with student participants to investigate preferences over inclusion of accident forgiveness in the insurance contract. The main question concerns how individual discount rates and risk attitudes influence the decision to purchase accident forgiveness. Both discount rates and risk attitudes were elicited through incentivized tasks. The author observes how safe drivers are in a driving simulator and places participants subsequently in driver risk categories, similar to how insurance contracts are based on recorded risk factors for drivers. The author concludes that both discount rates and insurance price reduce insurance purchases. Risk attitudes affect insurance decision-making only among those with a relatively high degree of risk aversion. We summarize the findings from driving simulator experiments that have studied choices in exogenous contexts in Table 2.

Table 2: Summary of findings from driving simulator experiments with exogenous context

Study Questions	Key Findings
	Virtual reality and driving simulators bridge internal validity of lab with external validity of field.
Route Choice	Risk attitudes elicited through lottery choices were similar to those elicited using a simulator. Risk attitudes are heterogeneous. Less exploration of risky routes occur when congestion risk is high
Safety	Subjective beliefs, experience and skill all influence decisions that affect crash propensities, controlling for risk attitudes Risk attitudes influence insurance decision-making

2.1.3 Laboratory Experiments without Driving Simulators

Information. Route Choice. Ben-Elia et al. (2008), Ben-Elia and Shiftan (2010), and Ben-Elia et al. (2013) analyse a pairwise route-choice experiment using induced monetary value for time and, with a large number of choice repetitions, investigated the effects of providing real-time information about travel time distributions on risk-taking behaviour. They gave their participants three types of pairwise route-choice tasks, with a Fast Route and a Slow Route in each pair. The routes also differed in riskiness. In the first type of task, “Safer-Fast”, the Fast Route was safer than the Slow Route. In the second type, “Risky-Fast”, the Fast Route was riskier than the Slow Route. In the third type, “Low-Risk”, both routes were relatively safe. Each participant was given 100 tasks for each of the three types, controlling for the order in which the types were played out. Each participant was also given an initial endowment and a cost of 2 cents per minute of travel time accrued throughout the experiment. Since the routes differed by 5 minutes in expected travel times, the expected cost of choosing the slower route was 10 cents in each task. The choice task was carried out in a two-route network. The authors randomly assigned the participant pool to two conditions: a control condition without any real-time information on travel times and a treatment group that did receive this information. The information given consisted of the minimum and maximum travel time for each route. The mean of this distribution was randomly selected before each drive, keeping both the difference in the means across the two routes and the range of both distributions fixed. Both groups were given feedback about the actual travel time after completing each period. The authors found that initial effects of information wore off with experience, and in both conditions behaviour became more stochastic over time in the tasks with higher risk for either of the routes.

Even though monetary incentives were used in this experiment, i.e. an induced monetary value for time at 2 cents per minute, the large number of repetitions and the speed at which each choice was carried out could have caused participants to use a relatively inattentive, mechanical choice process. Had fewer task repetitions been included, more attentive and deliberate choice process might have been chosen and the pattern of behaviour might have been quite different. Thus, while the incentives were salient in the sense that the monetary consequences were linked to their choices, the incentives might have not been enough to dominate other intrinsic values such as boredom and/or an urge to move quickly through the tasks might have confounded the incentives. Intrinsic value confounds are considered a violation of the dominance criterion in Induced Value Theory (see Harrison 1989).

Parking. Bergantino et al. (2015) provide a lab experiment on choice over parking alternatives. Student participants were presented with 128 scenarios in which they had to choose between three options that were either risky (the probabilities were known to the participants) or uncertain (the probabilities were not known to the participants). The risky option corresponds to situations where technology is used to provide information about parking availability, while the uncertain option has no such information. The payoffs for each option were in terms of an induced value of finding a parking spot (set at €10), the search time for finding a spot, and the walking time to the parking spot. Rather than induce the value of the search time in money, the cost was implemented so that participants had to stay in the lab after completing their tasks if they chose an option with higher search cost. Walking time was implemented as a walk to a separate room to get paid. Both the search time and the walking time varied across options. This design implied that the opportunity cost of choosing to search or walk was subjective and depended on the alternative activities that were available to participants at the time of the experiment. One of the 128 scenarios was randomly selected at the end to be the one that was actualized. Thus, only for that scenario was the risk or uncertainty played out to determine if they were paid the €10, and search and walking time were enforced. Responses were mostly in line with the predictions of standard economic theory, with the exception of walking time, which appears to have no significant effect on choices. Increases in parking cost and the cost of searching for parking reduced the demand for parking and increased the demand for walking. Technology enabled parking that transforms uncertainty into risk, increases the demand for parking.

Freight. Denant-Boemont and Hammiche (2010) conducted a lab experiment with student participants to test how agents react when one of two choice options is irreversible while the other is perfectly flexible and both have unreliable travel times. The motivation behind the design was the choices that haulers have between loading the truck on a train, and subsequently not being able to switch to a roadway option, or hauling on the highway with its more flexible route choices. In the experiment, participants made a freight choice in stage 1. If the choice was the highway, they had in stage 2 the opportunity to change to rail. No option to switch was available for rail. Payoffs to agents depended on tariffs for the road or rail option, and the risk of congestion with additional, costly travel time. The cost per minute of travel was induced and set to a constant that was the same for the two options. Travel times and the risk for congestion were calibrated to actual data from various studies on freight in France, thus giving the task a field flavour. Tariffs for the road and rail options varied across participants and could be low, medium or high. Further, there were two information conditions that varied across participants: in the Perfect-Information condition the state of the traffic in stage two was fully known to participants and could influence the route choice for those who selected the flexible option, and in the No-Information condition none of the participants knew anything about the state of congestion and travel times in stage two. Each participant made a sequence of 24 decisions, with feedback on earnings after each. Behaviour depended on the participants' risk attitude (value of travel time unreliability) and whether or not the perception of the costs and benefits of the options, as well as the probabilities of each possible state, matched their objective descriptions. The experiment showed that the results were in agreement with the theoretical model: Haulers' choices depend on the price for railways and on the information obtained by the hauler during their travel. However, the authors found evidence that individual behaviour tended to exhibit some inertia over multiple periods.

Safety. DeAngelo and Charness (2012) conducted a lab experiment with student participants on the choice of speeding when various mechanisms for enforcement are used. In particular, they focussed on the probability of getting caught and the penalty, when getting caught. The task was implemented only through text, no actual driving was involved. The rounds progressed at a rapid pace. All outcomes were induced in money terms. Participants were told that if they chose the option "not speed" they would get $0.60 while if they chose the option "speed" they would get $1.00. There were two separate experiments. In the first there was a 50% chance of being either in a case where the probability of being caught was 1/3 and the fine was $0.90 or in a case where the probability of being caught was 2/3 and the fine was $0.45. In either case the expected fine was $0.30. Participants made choices over 30 rounds. During the first ten rounds, the baseline, the only choice they made was whether to speed. After each round they got feedback on whether they were caught, and if so, what their fine was and their final earnings for the round. In the second set of ten rounds participants were asked to vote for one of the two enforcement mechanism at the beginning of each round, and the winning one was then implemented in that round. In a third set of ten rounds the participants were introduced to two new enforcement mechanisms, both with an expected fine of $0.50, i.e. higher than in the previous rounds. Again, the participants first voted for one of the mechanisms, and then made their decision to speed or not. In the second experiment the first ten rounds were conducted exactly as in experiment one. During the second set of ten rounds, instead of voting, one of the two enforcement mechanisms were imposed with certainty, but in the final set of ten rounds, voting was again introduced. The study shows that removing the extra layer of uncertainty after the first set of ten rounds increases the propensity to speed. The increased cost of speed in the last set of ten rounds in experiment one lead to less speeding. Further, when the regime shown to be less popular through voting was exogenously imposed, there was also less speeding.

The findings from laboratory experiments in exogenous contexts are summarized in Table 3.

Table 3: Summary of findings from laboratory experiments with exogenous context

Study Questions	Key Findings
	Money or time can be induced in the lab. There is an important role of information provision through technology.
Information	Initial effect of information wears off with experience Individuals are willing to pay more for information in situations with higher risks.
Parking	Induced search time as well as monetary costs Individuals react negatively to perceived increases in parking cost and the cost of searching for parking Technology enabled parking increases the propensity to search for more convenient parking rather than walking
Freight	Price and information influenced choice between truck and train Individual behaviour tended to exhibit some inertia over multiple periods
Safety	Individuals reduce speeding with increases in the expected costs Uncertainty in the enforcement regime results in reduction in speeding People are more likely to speed in a punishment regime that were implemented through popular vote

2.2 Choice in Endogenous Traffic Contexts

Endogenous traffic contexts are those where the risk of congestion or other outcomes depend on the choices made by the participants. The contexts are almost always lab settings for the simple reason that field experiments would generally have to be full field trials in order to generate endogenous system-wide outcomes.

2.2.1 Field Experiments

Route Choice and Pricing. An example of a full field trial is the Stockholm congestion charging trial conducted in 2006 (Eliasson et al. 2009). At great expense, the entry points to the inner city were outfitted with charging stations using cameras. Both entry and exit at these points were charged and the rates varied by time of day, with a minimum charge of 10 kr (about $1.35 in 2006) and a maximum of 20 kr. There was no exogenous variation of the charging scheme during the trial period. A reduction of 20-25% of traffic crossing the cordon was observed, and eventually a permanent charging scheme was implemented. The only outcome variable was the traffic volume going through the cordon. Individual vehicles or drivers were not tracked to observe any use of alternate routes (or modes). However, the authors report that experience of the field trial increased the acceptability of congestion pricing.

Full field trials such as this are both expensive to implement and also limited in the extent to which target variables can be varied, such as the pricing scheme. In a laboratory, costs are greatly reduced and a large number of exogenous variations in both charges and road characteristics can be implemented. There are trade-offs: one of the weaknesses of laboratory experiments is that the size of the group participating in any trial tends to be small, so that each individual participant has a direct and noticeable influence on the traffic outcomes. That said, it is well established in the laboratory that sometimes it takes surprisingly few participants to establish, for example, a situation of "many" (e.g., Huck et al. 2004). Table 4 summarizes this.

Table 4: Summary of findings from field experiments with endogenous context

Study Questions	Key Findings
	Cameras used at charging stations.
Route Choice and Pricing	Pricing reduces congestion Experience of pricing improves its acceptability

2.2.2 Laboratory experiments

The lab experiments with endogenous traffic conditions that we review here are undertaken under very stylized conditions where the only connection to transportation is in the labelling of the choice options and the outcomes. This type of experiment allows a longer time sequence in the lab with a larger number of observations in each traffic system over time. However, behaviourally it is possible that decisions differ in this type of setting compared to slower settings where each choice is carried out along with some real-time development of driving and traffic, such as in traffic simulations or field driving. These lab experiments use induced values for all benefits and costs associated with options and outcomes.

Information. Route Choice. Selten et al. (2007) reported an experiment conducted over 200 rounds where groups of 18 make a binary route choice in each round. Payoffs were presented in an experimental token currency and converted to a real currency (Deutsche Mark, using 2000-2002 exchange rates). In each round a driver received 40 tokens (about $0.24) from which the induced value of the actual travel time was deducted. One route (the Main Road) had more capacity than the other (the Side Road) and hence in any market equilibrium the predicted volume was higher (12 participants) on the Main Road was twice that on the Side Road, with token payoffs of 10 ($0.06) for each route. Between each round participants received feedback. In one condition they were told what their actual travel time was (as well as what their earnings were) and in another condition they were also given information about the travel times on the non-selected route. There were 6 sessions for each condition, thus 6 independent data points. The findings are similar to those of many experiments with multiple equilibria: there is a great deal of variation in participant choices with no strong indication of convergence. Based on these observations the study proposes two types of observed response modes: a direct response mode where a switch becomes more likely the lower the payoff was, and a contrary response mode where a switch becomes more likely the higher the payoff was.

Chmura and Pitz (2007) is an experiment with very extreme congestion payoffs: only participants who chose the non-congested route earned positive payouts. In transportation terms, such a payout scheme is like a late arrival penalty that exactly cancels the value of the trip and late arrival is a consequence of any congestion caused on the road with the majority of the traffic, no matter how small the majority is. Each group of 9 drivers repeated the same route choice task 100 times and were given feedback on whether the route they had selected was congested, and thus what their payoffs were for that period. As in Selten (2007) there was no actual driving so rounds progressed fairly quickly. Payoffs were denominated in experimental token currency and converted into Euros at the end of the experiment. Each token was worth 0.2€, with the maximum possible earnings being 20€. There were two information conditions: in one participants received feedback on whether the route they had chosen was selected by a minority or majority of the participants, and in the other they received additional feedback on the distribution of participants across the routes. There were 6 sessions for each condition, thus 6 independent data points. As in Selten (2007) there was considerable variation in the proportion of those who selected each route, but less so in the high information condition. In both treatments, participants who did not switch routes often earned significantly more.

Anderson et al. (2008) reports another lab experiment on route choice behaviour in small groups. Participants chose between a Safe Route with a fixed payout of $0.50 and a Risky Route where the payout depended on how many of the participants entered. The payoffs decreased by $0.50 for each entrant. All payoffs were in legal currency (US dollars). As in the Selten et al. (2007), and Chmura and Pitz (2004a, 2004b) the total traffic volume was close to the equilibrium prediction, but there was a significant amount of fluctuations across the routes. The authors also tested the impact of additional information on the number of participants who chose either of the routes, and they observed a significant reduction in fluctuations. They also found that tolls on the risky route reduced entry and thus congestion.

These studies were based on observing relatively small groups of participants where the route choice of each of them had a clear and noticeable effect on the congestion levels. Whether the findings are robust to larger groups where no individual driver has a direct influence on the congestion level remains an open question. It seems clear though that information provision has a stabilizing effect, consistent with conclusions drawn in Ben-Elia and Avineri (2015). This has implications for applications of information technology on congested routes.

Mode Choice. Denant-Boemont and Petiot (2003) investigates the valuation of information in a mode choice task. 3 groups of 10 participants made decisions over 10 rounds. Participants were informed about road capacities, after which they chose transport mode train or road. The train mode arrived on time with certainty. All players could purchase information about traffic levels on the road, i.e. how many participants selected the road. The purchase decision was private information – no other participants knew who or how many purchased information. Next, all players who chose the road selected their route, either the local or the highway. There was no toll on either route. The capacity on the local route was less than on the highway. They were again offered the chance of purchasing traffic information, this time for the conditions on the highway. Finally, participants chose departure time. Payoffs depended both on arrival time and travel time and both were induced in money. The data confirmed some theory predictions: participants' valued detailed information (about highway conditions) higher than general information (about roadway in the aggregate) and they bought information when the variance in payoffs was high, suggesting risk aversion. Despite assuming risk neutrality, their proposed theory provided a reasonable approximation for observed behaviour. Participants appear to be learning as the type of choices they make changes over time.

Networks. Mak et al. (2015) designed an experiment with a complex network consisting of a common origin, four different destinations with the same value, and multiple routes, each with three segments, many of which were shared between routes. The two conditions differed in the information given to participants. In the first, information about arrival times for various routes was only given after the route had been completed. In the second, information was given at each node between segments and participants made decisions about directions in a sequential manner. Five sessions of 18 participants each were assigned to each of the two treatments. The route choice task was repeated 50 times and no actual driving was involved. The authors report that there was no significant difference between the information conditions, with equilibrium predictions being approached rapidly in both conditions.

Pricing. Route Choice. Janusch et al. (2015) report a lab experiment of a two-route choice setting with varying tolls and six drivers in each group. One route was longer but with a certain travel time and the other route was shorter but with a risk of congestion. Value of time was induced in money and varied across participants. The experiment used student participants and there was no actual driving involved. In each session participants made route choices in three blocks of ten rounds each. The first block of route-choice rounds was a baseline without a toll. Before and after this first block there was a referendum over the acceptability of a toll. The second block featured an exogenously imposed toll, followed by another referendum, and the third implemented the outcome of one of the three referenda. Two conditions varied revenue allocations from the tolls and another two the information given before the vote. One revenue allocation treatment guaranteed that all participants were better off with the toll than without, the other one led to some participants being losers and some being winners. In one information treatment participants were shown the average total costs during the first two blocks of rounds for their group, and in the other they were shown each group member's costs. The authors report that the personal experience of tolling increases the acceptability of using tolls as a congestion management tool, and tolls generally reduce congestion.

Mode Choice. Chidambaran et al. (2014) report a lab experiment on mode choice using field participants in Hyderabad. Each session had 6 participants choosing between bus or car travel. The tasks did not involve any actual travelling. Participants in all 34 sessions first participated in 6 baseline rounds, followed by 6 rounds of either a bus subsidy (17 sessions) or a parking fee (17 sessions) and a final block of 6 rounds with a public coordination condition. In the public co-ordination condition one player would be randomly selected to be allowed to take the car in each round, and every participant was guaranteed to be selected exactly one time. Given the parameters used in the experiment, one participant taking a car and the other 5 taking the bus corresponds to the socially efficient solution. The user equilibrium was for 3 participants to take a car and the other 3 to take the bus. Recruitment selected only volunteers who had experience with both car and bus travel outside of the lab. The study finds modest effects due to parking fees and bus subsidies. The public coordination condition, although it increases earnings, is somewhat less effective at reducing car use.

Departure Time Choice. Gabuthy et al. (2006) tests the equilibrium predictions in a model of departure time choice with two routes that have different tolls and bottleneck capacities. There were 12 sessions of 8 drivers each making a sequence of 15 departure choices. The bottleneck capacity on one route is twice that of the other route. All values were induced in money and all participants were students. Feedback on travel time, arrival times, distribution of drivers, and payoffs are given after each round. There were two conditions: a low toll and a high toll on the high bottleneck capacity route. The low toll was found to be the most efficient. The equilibrium departure distribution was reportedly not reached but there was a trend towards the equilibrium.

Uncertainty. Mode Choice**.** Innocenti et al. (2013) report a lab experiment in mode choice. In each of three conditions participants chose between car and metro or between car and bus travel modes. 30 participants participated in the first condition, choosing between car and metro over 50 rounds. The metro option was described as a fixed, and known, total cost, consisting of a metro price and a monetized travel time cost. The car option had an uncertain cost due to endogenous congestion, determined partly by how many participants chose car over metro. In the other two conditions, with 15 and 17 participants each, the choice was between two uncertain options: car and bus, also for 50 rounds. The last two conditions differed in the fixed cost portion for the bus mode. Each condition was conducted only for one session. Payments were calculated in experimental token currency, converted to Euros during payments at the end of the session. The study finds that drivers do not appear to maximize profit, but display an intrinsic preference for car over metro and bus independently of the monetary incentives. The incentives at the margin appear very small, with an opportunity cost of less than 0.5 experimental token for taking the car instead of the metro or the bus under congested conditions, and each token being worth less than 10 cents. Given such modest incentives, it is not clear that the incentives satisfy the dominance criterion from Induced Value Theory.

Congestion cost. Denant-Boemont and Fortat (2013) conducted a lab experiment with student participants to test the effect of linear vs. exponential congestion cost, hypothesizing that the latter would lead to an enhanced reluctance of drivers to enter the congested activity. While the experiment was motivated by road congestion, no transportation labels were used in the tasks. Twenty sessions with small group size and exponential congestion were conducted, 10 with large group size and exponential congestion, 10 with small group size and linear congestion, and 5 with large group size and exponential congestion. The authors concluded that there were no differences in behaviour between an exponential and linear costs.

Departure time and bottlenecks. Schneider and Weimann (2004) evaluated the predictions of the single-bottleneck model proposed by Arnott et al. (1990, 1993). The experiment involved evaluating equilibrium predictions for departure-time choice, where the traffic system has a single bottleneck causing endogenous congestion. They conducted a set of two experiments referred to as the 'Magdeburg Experiment' and the 'Bochum Experiment' based on the location where the two experiments were conducted. The Magdeburg Experiment had three sessions, and in each session they had two groups of 20, which were later combined to form a group of 40. Therefore, there were 6 observations with a group size of 20 and 3 observations with a group size of 40. Each group played 30 rounds. The two treatments based on group size also differed in the parameters of the cost function. The authors found that in the 20 participant treatment the cumulative departure times observed were not statistically different from the equilibrium. However, in the 40 participant treatment they found the empirical distribution of the departure times was significantly different from both the equilibrium prediction and the Pareto-efficient outcome. Schneider and Weimann (2004) invoked payoff dominance as a possible explanation since the difference in expected payoffs between the equilibrium predictions and the Pareto-efficient outcome was relatively small. In the Bochum Experiment the same game was played with 10 players, and each player was given 10 vehicles to allocate across the departure times. They found that the observed departure time choices were close to Pareto-efficient distributions when the number of vehicles given to each participant was equal to the capacity of the bottleneck. They concluded that if each participant had control over multiple vehicles, the participants internalized the external effects of their choices by distributing the vehicles across alternatives, which results in the decisions tending towards Pareto-efficiency.

Ziegelmeyer et al. (2008) evaluated the impact of large group sizes and information on coordination and convergence to equilibrium. The experiment implemented Arnott et al.'s (1990) bottleneck model with departure time choices. Payoffs depend on induced value of time as well as early and later arrival penalties. They compare group sizes of 4 to group sizes of 16 and implement variations in public information. They found no evidence that either large or small group sizes had problems coordinating and neither information nor variations in delay cost had a significant effect. A potential explanation to the lack of information effects could be that information was about historic traffic and this was of only limited value since they were matched into new groups in each round.

Daniel et al. (2009) conducted an experiment to test a paradoxical bottleneck prediction. The experiment was carried out with 5 groups of 24 participants in two treatments, a high and low bottleneck capacity. Participants in each group made 50 departure time decisions, after 5 practice rounds. Travel times as well as early and later arrival penalties were induced in money. Each round progressed quickly since no actual driving was conducted. The consequences of all decisions were implemented by the computer as soon as all decisions were made. All participants had complete information about travel times of all drivers. Observations support the predicted paradox: as the capacity of one bottleneck is increased aggregate travel times increase causing a social loss. The authors also comment that behaviour is consistent with mixed strategy equilibrium play: average payoffs approach theory predictions, variability in payoffs across subjects decrease, and the distribution of departure times approximate the symmetric mixed-strategy equilibrium prediction.

Using a similar single bottleneck setup, Ramadurai and Ukkusuri (2007) studied departure-time choice while accounting for the dynamics of traffic with homogeneous travellers having the same preferred arrival times. They theoretically showed the nonexistence of a Dynamic User Equilibrium in a discrete time single-bottleneck model. To test their theory, they designed a multiplayer online experimental network game to simulate traffic with a single bottleneck. The dynamics of traffic was simulated using the cell-transmission model. They paid the participant that had accumulated the most points at the end of the experiment. This tournament style reward system tends to increase risk taking among participants. They found support for nonexistence of equilibrium, and also observed less variability in choices when information was provided.

Paradoxes. Paradoxes in modelling transportation systems have played an important role when arguing in favour of informing policy decisions with transportation models. Researchers have used EE to study paradoxes associated with mode choice, route choice and information.

Downs-Thomson. The Downs-Thomson Paradox states that additional road capacity causes an increase in total travel cost for users that choose between private car and public transit. Recent experimental work (Denant-Boemont and Hammiche 2012; Hartman 2012; Dechenaux et al. 2014) has confirmed the existence of the Downs-Thomson Paradox. While all these studies found the aggregate choices to be close to the equilibrium, they also found systematic differences in individual level choices and variability. Denant-Boemont and Hammiche (2012) used experimental groups of 15 subjects, where 14 were travellers choosing between road and transit, and the 15th subjects was the transit operator choosing transit capacity. The experiment was executed under two conditions: high and low road capacity. The transit operator is predicted to choose the same capacity, the minimum one, in both of these conditions. In eight sessions subjects initially experienced the low capacity condition and subsequently the high capacity condition, repeated over 20 rounds. In another eight sessions the order of the low and high condition was reversed. Subjects were paid for 4 randomly selected decisions from the 40 that they made. As a design consideration, this feature allows subjects to reduce the risk, compared to when only one decision is selected for payment, since they could form risk reducing portfolios. The study reports that transit operators are not affected by the road capacity treatments, as predicted, but that they invest at a higher level than suggested by theory, although this tendency decreases over repetitions. Travelers respond to the exogenous shift in capacity as expected. The Downs-Thomson paradox is confirmed for the sessions that experience an increase in capacity, but the opposite effect is not found for the sessions that experience a decrease in capacity. In Hartmann (2007, 2012) participants are given a choice of an uncongested highway or a potentially congested bridge, with induced value of time. The travel time on the bridge is a known linear function of the number of travellers. In one experiment three conditions are used that vary in the distribution of the induced value of time. In another experiment, one condition was included where value of time was "home-grown" rather than induced in money. In this condition participants have to wait in the laboratory for the amount of time that they have been delayed due to congestion. The study shows that increased heterogeneity in value of time leads to less variation in the proportions choosing each of the two options and to faster learning. The study also shows that the Downs-Thomson paradox can be avoided through tolling the shorter route. Decheneaux et al (2014) test the paradox in a setting where travel costs due to congestion changes in a continuous fashion for the road option but in a discrete fashion for the public transit option. They conduct an experiment under four conditions, varying group sizes (10 and 16), public transit pricing structures (average cost and constant cost), and varying road capacity. Each subject participated in two conditions, and the order of the conditions varied across sessions. Aggregate choice behaviour changes according to predictions across the conditions, although by less than predicted, but there is a great deal of variation at the individual level. The data generally supports the Downs-Thomson paradox.

Braess Paradox. The Braess Paradox states that an additional link (or capacity) to traffic networks leads to deterioration of the overall traffic for certain networks. This prediction is driven by the assumption of myopic and self-interested agents trying to maximize their own utility which leads to an equilibrium which increases social cost compared to the social optimum. Rapoport et al. (2006) studied the Braess Paradox with three levels of demand and found confirmatory evidence for the theory. Specifically, at the lowest level of demand, the addition of a Braess link improved the efficiency of the network, but with larger demand an additional link led to a deterioration of the overall traffic of the network. The experiment was carried out using 4 groups of 10, 2 groups of 20 and a single group of 40. In each group route choice decisions were made 40 times. Following up on this study, Rapoport et al. (2009) studied the Braess Paradox with larger groups of participants and found that the paradox does indeed exist when group sizes are larger. Three groups each of 18 subjects participated in either a network where a link was added or a link was removed. In both of these cases is the paradox confirmed: subjects choices converge to the lower payoff equilibrium. In a second experiment 6 groups of 18 subjects participated in route choice decisions in a larger network with more complex choice considerations. At the aggregate level the data supports the equilibrium paradoxical predictions. Both Rapoport et al. (2006) and Rapoport et al. (2009) show that at the individual level behaviour does not conform with equilibrium predictions, even though it does at the aggregate level.

Meinhold and Pickhardt (2009) studied pricing in a Braess network and theoretically derived the maximum price that could be charged without causing a Braess paradox. The experiment was conducted with paper and pen in groups of six. All groups first experienced the network without pricing over six rounds. In five groups subjects then experienced six rounds of pricing that was predicted to generate the Braess paradox followed by six rounds of pricing that was the predicted maximum not leading to Braess paradox. Another five groups experienced the pricing in the opposite order. Behaviour is close, but not equal to, the Nash Equilibrium predictions. Excess burdens are higher in the Braess paradox

Morgan et al. (2009) perform experimental tests of both the Downs-Thomson and the Braess paradox in very similar networks. Groups of eight subjects participated either in a Downs-Thomson or a Braess paradox and for each there were separate groups experiencing the changing conditions in opposite orders. Each subject made 120 decisions, 60 in each condition. The conditions for the Downs-Thomson paradox was a change in capacity and the conditions for the Braess paradox the existence of the extra link. In both of these experimental conditions behaviour is close to, but not equal to, the equilibrium. An important policy conclusion from this paper is that planners can effectively improve travel times by improving those routes least sensitive to network congestion (i.e., "least congestible route"). On the contrary, an increase in demand (i.e., "size") can make the addition of a link counterproductive for society.

<u>Information paradox.</u> Rapoport et al. (2014) test another paradox introduced by Lindsey et al. (2013). According to this paradox, under certain conditions pre-trip information about travel conditions can lead to longer travel times. The experimental implementation has two types of travel conditions for each route, a high and a low capacity. There are two treatments: one where the conditions on the two routes are correlated and one where they are not. Five groups of 20 participants were assigned to each of the treatments. In each treatment participants first made

route choices over 80 rounds without prior information on travel conditions, and then over 80 rounds with prior information. The authors find that information is beneficial when route conditions are uncorrelated but not when they are perfectly correlated.

The ability to control conditions and generate high internal validity in the lab has resulted in a great variety of questions being asked and theories being tested. It has become clear that variation in experimental designs and protocols affect the findings. Table 5 summarizes the findings from endogenous laboratory experiments. In the next section we attempt to summarize the insights gained from the experimental applications.

Table 5: Summary of findings from laboratory experiments with endogenous context

Study Questions	**Key Findings**
Information	Significant variation in individual route choice but information reduces such variability Valuation of detailed information is higher than of general information
Pricing	Pricing has a modest influences on mode choices but reduce congestion in route choices; more so when prices are lower. Personal experience of pricing influences acceptability of pricing
Uncertainty	Intrinsic preference for car over transit despite greater uncertainty
Congestion cost	No significant behavioural difference between a linear and an exponential congestion costs
Departure time choice	Mixed findings on effect of group size Support for mixed strategy Nash Equilibrium, especially when individuals control multiple vehicles.
Paradoxes	Confirmation of theory predictions: Increase in capacity can result in increased congestion Information sometimes result in deterioration in traffic systems

3. INSIGHTS INTO TRANSPORTATION BEHAVIOUR FROM EXPERIMENTS

Behavioural experiments in transportation science have been driven by three main needs:

a) Testing mechanisms and technologies to evaluate their effectiveness (eg. incentives and monitoring to influence behaviour)

b) Testing theories and the fundamental assumptions they are based on (eg. testing paradoxical predictions and equilibrium properties)
c) Exploring relationships between key characteristics of contexts or participants (eg. impact of group size on choice behaviour, impact of various types of information on choice behaviour, and the impact of subjective beliefs on unsafe behaviour and the propensity to crash)

Key findings from our review of applications of EE to transportation are shown in Tables 1-5. As stated earlier, the use of actual consequences in EE significantly reduces hypothetical bias that exist in many stated preference surveys.[6] Furthermore, EE has contributed significantly towards beginning to understand choice behaviour in both exogenous and endogenous transportation systems. While our review demonstrates that quite a few transportation experiments have been undertaken, the applications and designs vary such that not many common insights have yet been produced. The state of affairs is therefore one in need of replications. Some experiments also used questionable incentive mechanisms, and the reported findings would need to be reproduced under more salient incentives. Further, the use of driving simulators for endogenous behaviour has not been undertaken as yet.

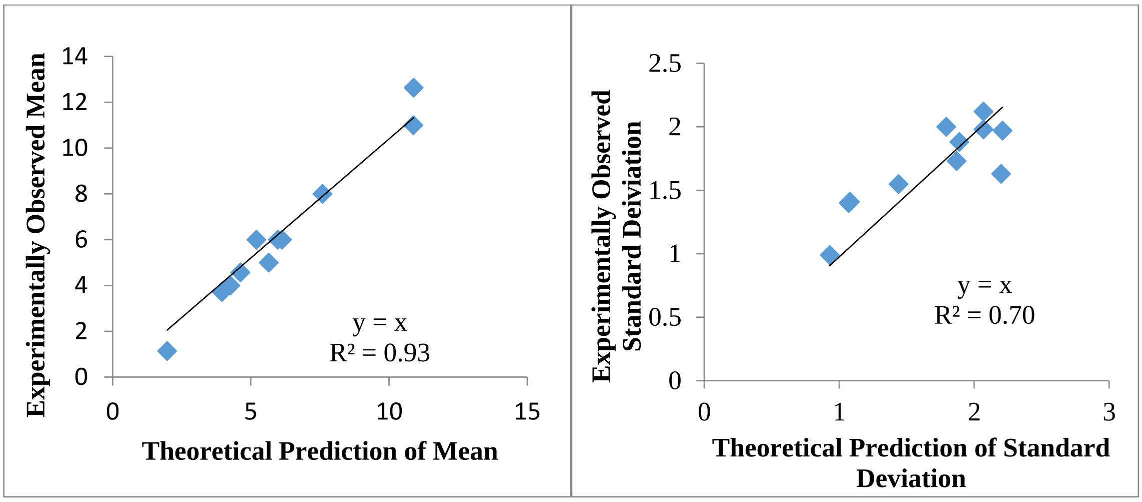

FIGURE 1: Comparison of Mixed Strategy Nash Equilibria predictions for binary choices in endogenous transportation systems with experimental data from (Dechenaux et al. 2013; Selten et al. 2007; Hartman 2012; Morgan et al. 2009; Rapoport et al. 2009). Data shows proportions of one of the binary choice options.

Using pricing or information to influence behaviour has been shown to have some success. Effects of manipulations sometimes wear off over time, and at other times the pure effects of being in an experimental situation creates changes in behaviour. There are, however, many details that differ across experiments and the detailed results are affected by such differences. In endogenous contexts there is general support for equilibrium predictions at aggregate levels, but at the individual participant level there is both variability and heterogeneity in responses.

[6] Reductions in the hypothetical bias of stated choice surveys have also been achieved without adding salient incentives. See Hensher (2010) and Fifer, Rose and Greaves (2014).

Thus, while at the group level a single market equilibrium may be observed, at the individual level there is a strong tendency toward stochastic choice or mixed strategy Nash Equilibrium behaviour.

Transportation modellers have relied heavily on equilibrium theories for planning purposes. Theories, such as the Wardrop Equilibrium (Wardrop 1952), mixed-strategy Nash, and Stochastic-User Equilibria (Daganzo and Sheffi 1977), have been investigated in the literature. The replications of theoretical paradoxes in the lab provide some confidence in these theories. Almost all studies find significant switching at the individual level, suggesting either mixed-strategy Nash, or Stochastic-User equilibria. In Figure 1 we compare the mixed-strategy theoretical predictions of mean and standard deviations with those observed experimentally in the literature (Dechenaux et al. 2013; Selten et al. 2007; Hartman 2012; Morgan et al. 2009; Rapoport et al. 2009).

As is demonstrated in the R^2 fits of the y=x lines in Figure 1, the theoretical mixed-strategy equilibrium predictions agree well with the experimental mean behaviour findings, even when looking at the standard deviations. However, underlying these mean behaviours is a great deal of individual variations and understanding these is critical to understanding factors affecting reliability in the transport system (travel time reliability) and how people respond to it. Two distinct behavioural response modes were found in route switching behaviour by Selten et al. (2007): *the direct-response mode* and *the contrary-response mode*. Further, Dixit and Denant-Boemont (2014) showed that, although a mixed-strategy equilibrium performs reasonably well when there are multiple equilibria, a Stochastic User Equilibrium -- a special case of Quantal Response Equilibrium (Mc Kelvey and Palfrey 1995, 1998; Goeree et al. 2008) -- provided much better predictions.

4. DESIGN OF BEHAVIOURAL EXPERIMENTS ON TRANSPORTATION

The methods from EE provide tools to implement behavioural control over participants when testing theories and policies. What distinguishes EE from Stated Choice experiments is the reliance on Induced Value Theory (Smith 1976, 1982). By inducing preferences (valuations) and/or beliefs following the conditions stated in Induced Value Theory, the experimenter gains control over many unobservable factors that influence choices. This is necessary for internal validity, i.e. for matching the experimental conditions to the assumptions of theory. Smith (1976) argued that the following three conditions are required to properly control participants' innate characteristics and preferences:

1. Monotonicity: This condition points to the nature of the *incentives* used. They should be such that experimental participants must prefer more of the reward medium over less, thus never being satiated within the experimental task. In the context of transportation systems, agents make choices in the *money* and the *time* domain. Preferences over money and time in transportation decisions outside the experimental laboratory can be assumed to be monotonic and the incentives used in transportation experiments should reflect this.

2. Salience: Given the institutional rules, and participants' understanding of the link between actions and payoffs, the rewards received by a participant should depend on their actions (as well as that of other participants, in experiments involving interactions).
3. Dominance: Changes in participants' utility from the experiment should come predominantly from the reward medium, with other influences being negligible. This condition is the most problematic of the three conditions, since "other influences" may not be observable by the experimenter. Dominance becomes more plausible if the salient reward is large; this proposition itself is eminently testable in the laboratory. In addition, this condition is relevant for the no-deception norm widely accepted in EE, as participants second-guessing whether the experimenters will actually pay them is likely to lead to loss of dominance of the induced incentives.

In essence, Induced Value Theory postulates that the incentives provided in this manner can be treated as the only argument in experimental participants' utility functions.

Due to the heavy reliance on these methods to ensure experimental control, it is critical to understand the strengths and pitfalls of various incentive mechanisms with respect to payment protocols, salience, and dominance.

4.1 Incentives and payment protocols

The literature offers various payment protocols ranging from lab money and actual money to randomized payments. We critically review these methods and identify their inferential limits.

4.1.1 Payment Units

A common incentive mechanism used in laboratories are *experimental monetary units* (referred to as Experimental Currency Units or ECU by Denant-Boemont and Petiot 2003; Lab Money or LM by Rapoport et al. 2007; Points by Denant-Boemont and Fortat 2013, Gabuthy et al. 2006, and Hartman 2012; Taler by Selten et al. 2007; Time by Dixit et al. 2013) which are converted to money at the end of the experiment, with the exchange rate typically – but not always – known by participants. Such experimental monies are typically used to study game-theoretic interactions to test equilibrium predictions as well as paradoxes in laboratory settings.

Bohm (2002) pointed out that the use of experimental monies has its roots in voluntary public exchange games, such as public good games, and was intended to contribute to *context-free experimental design*. There are examples of context-free experiments in the transportation literature (for example, Denant-Boemont and Fortat 2013; Denant-Boemont and Hammiche 2013; etc.). Context-free experiments have been criticized by several authors (Binmore 1999; Loewenstein 1999; Harrison and Rutström 2001; Bohm 2002) who believe that participants tend to create a context of their own if it is not provided for them in the experimental setting, resulting in a loss of control. In transportation experiments, the problem of field referents is easy to overcome by conducting experiments in transportation contexts; as Bohm (2002) points out, some such framing will increase external validity and "*contribute to meeting Vernon*

Smith's (1982) parallelism precept" (p. 124). The parallelism precept states that the behaviour observed in the laboratory applies to non-laboratory conditions where similar *ceteris paribus* conditions apply (Smith 1982, p. 936).

Whether choice of the reward medium and/or labelling (ECU, LM, points, tokens, or time) play an important role in the experimental setting remains an open question. Transportation choices are predominantly made over time and money, and introducing labels different from time and money, such as tokens or lab money, could trigger loss of control and loss of external validity. We consider the question of labelling a design question that is in need of further exploration. Based on the results of earlier studies, we advise payment protocols and incentives that use labels consistent with the transportation context.

4.1.2 Cost Functions and Virtual Reality

From a traffic perspective saliency is achieved by using costs and tasks that are similar to those in the field. Denant-Boemont and Fortat (2013) studied the impact on equilibrium behaviour due to non-linear cost functions, which are similar to the travel time functions (Bureau of Public Roads, 1964) used in transport models. Good examples of protocols to study traffic interactions that were salient in the laboratory are those by Ramadurai and Ukkusuri (2007) and Andersen et al. (2011). In these studies travel costs were generated from traffic simulations that were consistent with traffic-flow theory and traffic-flow phenomena. Specifically, Ramadurai and Ukkusuri (2007) used a cell- transmission model while Andersen et al. (2011) used traffic micro-simulation software. A much higher degree of external validity is achieved when the choice involves the actual driving task, such as in the use of virtual reality driving simulators, which have been used for route-choice studies (Andersen et al. 2011 and 2014; Dixit et al. 2015) and safety studies (Dixit et al. 2014).

The ability for Virtual Reality (VR) to create the experience of "presence", i.e. the sense of being in the Virtual Environment and feeling as if you are in reality, is what makes it powerful. However, generating high quality scenarios with "presence" using virtual reality can be costly. Further, poorly designed virtual environments can lead to scenario rejection (as explained in Fiore et al. 2009). Sanchez-Vives and Slater (2005) identified four main ways of evaluating "presence": (a) Questionnaires about experience which have been found to be unstable. (b) Replication of behaviour in the virtual environment (Yan et al. 2008) (c) Replication of physiological responses in the virtual environment (d) Counting number of breaks in presence in the virtual environment, also referred to as Breaks in Presence (BIPS).

Further, Sanchez-Vives and Slater (2005) in their review found that the following five factors in the virtual environment provide realism: (a) The reported presence is found to be higher with higher graphics frame-rate, and a minimal critical frame-rate of 15 Hz should be maintained. (b)Realistic visual cues with dynamic lighting shadows and shapes (c) Realistic audio cues (d) Realistic haptic cues (e) Realistic body engagement as in being involved in the driving task. (f) Realistic virtual body representation involves being able to see their own bodies for example in a head mounted display.

In the field of transportation driving simulators have been extensively used to introduce a sense of immersion among the participants that have virtual and audio cues from the field. In addition to the natural-like cues, the participants are also engaged in the task of driving, which brings it closer to how choices are made in the field. Hence there seems to be a continuum of experiments that try to achieve external validity through incorporating virtual reality.

4.1.3 Effects of Repeated Rounds

Many transportation experiments have focussed on studying group interactions based on game-theoretic predictions (equilibrium and paradoxes); they utilise multiple repetitions, or rounds, of the same game, occasionally called Groundhog-day repetitions (Hertwig and Ortmann 2001). Bohm (2002) identified this as a "*useful approach when analysing the properties of real world markets*" where agents are also typically involved in repeated transactions. Transportation is one such market, where people interact and on a daily basis decide whether to make a make a trip, which route to choose, and when to depart. When conducting experiments with multiple rounds to study transportation phenomena, we need to control for:

I. Learning

Game-theoretic experiments have shown strong evidence of learning behaviour in the laboratory; see for example Denant-Boemont and Petiot 2003; Chmura and Pitz 2007; Schneider and Weimann 2004; Klügl and Bazzan 2004; Selten et al. 2007; Ben-Elia et al. 2008; Denant-Boemont and Hammiche 2012. Most experimenters have found strong evidence for reinforced-learning models, or variants of it such as the payoff-sum model (Selten et al. 2007). There is also evidence of Bayesian learning in the presence of information (e.g., Denant-Boemont and Petiot 2003). To test equilibrium predictions, it is critical to recognize the learning process and to allow for a sufficient number of periods to pass before the experimental data is used to test theoretical predictions. Ben-Elia et al. (2013) used the mean standard deviation of choices as well as the scale parameter of the logit choice models to evaluate the learning effects.

Experimenters also need to recognize the potential for end-of-game effects: changes in behaviour as the round repetitions near their end. Such effects have been found in many experiments (Roth et al. 1988; Gneezy et al. 2003) and may be attributable, at least partly, to more risk-seeking behaviour towards the end. No such behaviour has been reported in transportation experiments to date.

II. Wealth Effects

A related matter is the potential for behavioural changes due to wealth accumulation during an experiment. There is evidence from individual choice experiments that wealth accumulation influences the amount of risk participants' are willing to take (Holt 1986; Andersen et al. 2012; Cox et al. 2015), usually reducing risk aversion. In repeated stage games where participants are paid for every choice, their accumulated wealth would typically increase, possibly resulting in decreases in the aversion to risk,. In the traffic context, one important source for wealth accumulation is the time savings consequences of the choices, since the value of time is usually induced with money. Dixit et al. (2015) provide evidence of wealth effects in a route choice experiment.

III. Random payments

One popular way of avoiding wealth effects is to only pay for one of the experimental rounds, randomly selected. Several carefully designed experiments provide reassuring evidence in favour of the random selection method for individual choice experiments (Starmer and Sugden 1991; Cubitt et. al. 1998; Hey and Lee 2005). We are unaware of studies that test the validity of the random-payment protocol in game-theory experiments, especially in the context of traffic. Aside from avoiding wealth and portfolio effects that may emerge if participants are paid for each task, it has been suggested that the random-payment protocol may help to economize on the cost of data collection (Davis and Holt, 1993). We are not persuaded by that argument, as participants are typically brought to the lab with expectations about average earnings.

Random-payment protocols are not conducive to test non-expected utility models such as rank-dependent utility, prospect theory or any theory where there is probability weighting (Cox et al. 2015; Harrison and Swarthout 2014), as they create situations of compound lotteries, wherein the probabilities do not aggregate perfectly, violating the independence axiom of expected utility theory. Therefore, any analysis of data collected through a random- payment protocol is limited by its reliance on expected utility theory (or other, less common, assumptions).[7]

TABLE 6: Comparison of payment protocols for participants and tasks

Pay Task / **Pay Subject**	**One (Best)**	**One (Random)**	**All**
One (Best)	Non-Salient	Non-Salient	Non-Salient
One (Random)	No Wealth Effects, Data Truncation	Only EUT, No Wealth Effects	EUT, Non-EUT, Wealth Effects
All	No Wealth Effects, Data Truncation	Only EUT, No Wealth Effects	EUT, Non-EUT, Wealth Effects

To economize on the costs associated with paying participants, researchers such as Ramadurai and Ukkusuri (2007) applied a scoring rule that had the top performers in the game earn rewards. However, this approach is not salient or externally valid since it changes the task into a tournament, unlike driving situations external to the lab. In addition, such tournament payment schemes are riskier and thus may lead to less risk taking behaviour. One way out, arguably, is to pay participants for their best performing task. When paying for best performance, data subsequent to a task where the participant makes the maximum possible is uninformative, since beyond that there are obviously

[7] Many non-expected utility theories allow violations of the Independence Axiom. Applications of random payment protocols rely on this axiom for its incentive compatibility property. However, Holt (1986), and Harrison and Swarthout (2014) discuss alternatives to this axiom that would maintain incentive compatibility of the random payment protocol even in some non-EU applications.

no further performance incentives. This can then lead to limited usefulness of the data. The issues are summarized in Table 6. Caution is also advised to make sure that the incentives, whether paid randomly or not, are sufficiently large to dominate other factors that may influence the choices.

4.1.4 Money vs. Time Incentives in the Laboratory

Few studies have explored utilizing time instead of money as an instrument to induce value. Bergantino et al. (2015) studied the trade-off between the cost of parking, time spent searching, and time spent walking from the parking. The experiment required participants to wait in a separate room for a time period that reflected the time spent searching for parking in the experiment. The participants were then asked to walk different distances to receive payments based on the "walking time" determined in the experiment. Kemel and Paraschiv (2013) and Abdellaoui and Kemel (2013) have also studied the trade-off over time and money but do not discuss specific applications to transport.

Hartman (2012) also explored the implications of using *time* to induce value in route choice experiments. Like Bergantino et al. (2015) he also asked the participants to wait for a period of time that was determined by the outcomes of the experiment. He found that very few participants gave up money to wait less, and these participants indicated that they were constrained in time due to other commitments. This finding has two important implications:

- The decision by the participant to participate entails them making a trade-off between the expected payoff from participating in the experiment and the expected duration. The fact that they participate in the experiment reveals that the expected payoff from the experiment is larger than the opportunity cost of the time committed to the experiment. It is possible, therefore, that the value of the extra time is fairly small in such laboratory environments.
- Hartman's finding that participants who indicate that they have time constraints paid money to wait less, implies that the value of time varies with circumstances. Thus, a participant who has cleared enough time to finish the experiment will have a lower value of time than a participant who has many alternative options that compete for the time allocated to the experiment. Further, it is unlikely that either of these measures of the value of time are appropriate in applications to driving choices. Such variations in the value of the reward medium is much less likely if the reward medium is money.

4.1.5 Designing Groups

In interactive traffic experiments, the incentives not only depend on one's actions, but also on the actions of other participants in the group. This makes it critical to determine how the participants are assigned to groups and their computer terminals. Gabuthy et al. (2006, p. 279) "*randomly assigned to visually isolated computers*", and then ensured that the participants interacted repeatedly within the same group across periods. This matching mechanism is usually referred to as a "Partner design", notwithstanding the fact that participants typically do not know who they are match with (although they do know that they are matched with the same participants). Some such matching mechanism is critical since in the field the *same strangers* often interact on a daily basis in making decisions and they do learn about their aggregate behaviour. A matching mechanism that changes the participants of the interactive traffic experiment every period –called "Stranger design" – is appropriate for environments where the actors are not the same and there is not much opportunity to learn about how others behave. Ziegelmeyer et al. (2008) used such a "Stranger design": participants were randomly assigned in each period to evaluate a departure time choice. The authors found no learning and no impact of information. To test theories that assume that actors interact only once, it may be necessary to use a "Perfect Stranger design", where it is guaranteed that no participant will ever be in a group with any of the same other participants. The choice of group matching mechanism must be carefully determined based on the research question in mind.

4.2 Salience

Experiments are considered salient when the experimental rewards are linked to the actions of the participants in a way consistent with the predictive theory, and when the participants have a clear understanding of this link. This implies that use of flat incentives to participate or provide responses cannot be considered as a salient experimental design since other factors will then influence the choices in unobservable ways.

4.2.1 Laboratory

Apart from ensuring that actions are incentive compatible, it is critical that the participants' have a clear understanding of the link between their actions and rewards. As discussed earlier, labels are an important technique used to introduce context in laboratory experiments. However, to ensure a salient design special care should be taken that the properties of the labels are actually reflected in the experimental design. For instance, it is well known that the distribution of travel time observed in the field has a unique peak in the probability distribution (Brilon et al. 2005), and if travel time is used as a label in a laboratory experiment, then the underlying distribution should have the same property. This issue was observed in the work by Innocenti et al. (2013). The authors claim that they used "*explicit labels of travel modes*" (p. 159) in their lab experiment, inducing participants "*to take into the lab the preferences applied in real life, and to rely on heuristic rules to determine their choice*" (p. 159). The properties of the distribution of the travel times imposed in their design was, however, not consistent with those expected in the field. This is shown in Figure 2. This could lead to scenario rejection, a type of behaviour that is unrelated to the research focus and simply a response to the unrealistic scenario presented.

Figure 2 shows that the probability distribution of the travel costs for the car (reproduced from Table 3, p. 161 in Innocenti et al. 2013) is decreasing in both of their experimental treatments (metro and bus), while a unimodal distribution would have been expected based on field experience. Therefore, the "*car stickiness*" that the authors find might result from "field referents" (Harrison and Rutström 2001) which were not properly captured in the experimental design. Participants who bring beliefs based on field experience into the lab could have made decisions based on a subjective uni-modal distribution.

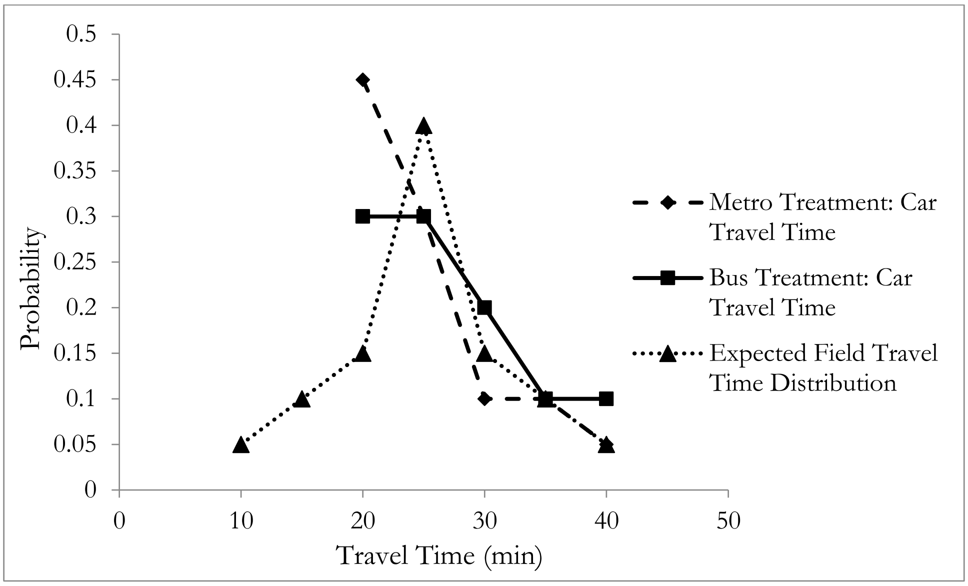

FIGURE 2: Comparison of probability distribution provided in the Innocenti et al. (2013) experiment compared with that expected in the field

4.2.2 Field

Field experiments in EE commonly use participant pools that represent the target population of the policy question investigated. Dixit et al. (2015) compared adult commuters (field participants) and students (lab participants) from different regions with respect to risk attitudes inferred from route choices in driving simulators and found that with minimal experience they behaved on average similarly. Use of field participants instead of students helps provide an added level of external validity. Andersen et al. (2010) show that the only difference in risk attitudes between an adult field population and students in Denmark is due to the more limited demographic distribution among students.

Maintaining a salient design in the field is more challenging due to other strategic factors. For example, Dufwenberg and Harrison (2008), celebrating the contribution of Peter Bohm to field experiments, describe an experiment he undertook in 1980 for a local government in Stockholm and in which he considered expanding a bus route to a major hospital and a factory. The experiment involved eliciting people's willingness to pay for additional service on the bus route. However, no one participated in this experiment. As identified by Dufwenberg and Harrison (2008), the participant pool had reportedly held a meeting of their own and had decided

"*(1) that they did not accept the local government's decision not to provide them with regular bus service on regular terms;*

(2) that they did not accept the idea of having to pay in a way that differs from the way that "everybody else" pays (bus service is subsidized in the area)—the implication being that they would rather go without this bus service, even if their members felt it would be worth the costs;

(3) that they would not like to help in realizing an arrangement that might reduce the level of public services provided free or at low costs. It was argued that such an arrangement, if accepted here, could spread to other parts of the public sector; and

(4) on these grounds, they advised their union members to abstain from participating in the project."

This case highlights the difficulties of conducting field experiments where participants perceive a strategic value to their options outside the experiment being affected by their actions in the experiment. This makes it imperative to carefully understand and design field experiments. More recently, Stockholm conducted a field experiment on congestion pricing that helped evaluate its efficiency and apparently changed public opinion (Eliasson et al. 2009), demonstrating how field experiments may change not only the opinion of policy makers but also those of other stake holders. Similarly, Nielsen and Jovicic (2003) conducted an earlier AKTA road pricing experiment in Copenhagen using GPS trackers to record route decisions. One concern with in-vehicle recording systems is the fear of surveillance, but Nielsen and Jovicic (2003) report that no such fear was found in their sample. Similar lack of fear findings were observed by the more recent study by Andersen et al. (2011) as well as the Pay-As-You-Speed literature (Agerholm et al. 2008; Hultkrantz and Lindberg 2011; Bolderdijk et al. 2011).

Recently, incentive-based mechanisms have become increasingly popular as a tool towards promoting off-peak travel to decongest roads (Merugu et al. 2009) and public transit (Pluntke et al. 2013). The basic idea of these mechanisms is that people are incentivized with points for favourable behaviour (walking, travel during off-peak, etc.). The INSINC program in Pluntke et al. (2013) exemplifies a number of factors where caution in design needs to be applied in order to avoid participant strategic behaviours leading to increased program costs. For example, the additional rewards for off peak travel is likely to attract more off peak commuters than on peak commuters to participate in the program, particularly since the details of how credits are given can easily be transmitted with the social recruitment mechanism (participants get bonus credits if they recruit friends and families to participate). In addition, participants that do not shift to off peak or that are not particularly active in attracting new members are given special offers to add incentives. This type of incentive, if not implemented cautiously, can also lead to strategically avoiding shifting to off peak or not inviting friends in anticipation of receiving such personalized bonuses.

4.3 Dominance

Dominance is critical to ensure control. For example, participants often care about the rewards earned by other participants, for example through guilt when earning more than others or envy when earning less, which might lead to uncontrolled influences on behaviour. If the experimental procedures make it impossible to know or estimate others' actual rewards (Smith calls this *privacy*) then this issue is successfully addressed. Another component such as *demand* *effects*, arising from participants' desires to help (or hinder) the experimenter, again leading to uncontrolled influences on behaviour. For example, when using labels on the actions, such as traffic congestion, participants may want to help the experimenter reach an outcome of less congestion independently of any induced value preferences. While this may seem a desirable effect on first glance, it is the fact that we cannot observe the expressed preferences that make it hard, or even impossible, to understand the causation involved. Avoiding labels will mitigate such demand effects.

Ben-Elia et al. (2013) found behaviour was consistent with the presence of fatigue in the treatment group where no information was provided, but not in the informed treatment group. This is consistent with low monetary consequences that may be counterbalanced by the presence of information that keeps the subjects engaged.

To achieve dominance, it is also important that participants do not suspect that the task instructions are deceptive. In some experimental applications, outside of economics, deception is often used. If economics experimenters recruit from the same participant pool as researchers who use deception, there is a loss of control over dominance since the participants may not trust the instructions of how actions lead to earnings.

Schneider and Weimann (2004) present the only study in transport in which the impact of dominance on the results is discussed. They stated that "the payoff function is flat. Therefore, large deviations from the Nash distribution of departures have only a small impact on the payoffs." (p. 144) *Pay-off dominance* was originally discussed by Harrison (1989) in response to findings by Cox et al. (1982, 1983, 1985, 1988). Based on this argument, the authors refused to reject the Nash equilibrium prediction based on action choices, and uses instead the stability of the observed patterns across cohorts as evidence that the "participants behave 'as if' they are in Nash Equilibrium." (Schneider and Weimann 2004; p. 144). Their study directs attention to the importance of having payoffs that dominate by being high enough to drown out other influences. This makes it even more critical to carefully consider the nature of dominance of other incentives and payment protocols.

4.4 External Validity

Internal validity implies that the experiment matches the assumptions of the theory with strict control of any confounding influences. Maximizing internal validity leads to very artificial and context free tasks. External validity, on the other hand, implies a precise matching of the contexts and cues from the intended policy application. Maximizing external validity leads to having to relax many of the controls required for internal validity. Three main types of experimental setups can be identified (see Figure 3): 1) Laboratory experiments 2) Virtual reality or driving simulator experiments 3) Field experiments. At one end of the spectrum lies the laboratory experiment that, if carefully designed, can provide a very high degree of internal validity. The structures, information conditions, and preferences can be made to match the theory very precisely. At the other end of the spectrum lies the field experiment that can provide a very high degree of external validity (i.e. predictions that are valid for the intended policy context outside the experiment). Strengthening external validity leads to a weakening of internal validity. For example, use of labels on actions and outcomes that have field relevance in lab experiments leads to a loss of control over the preferences that participants apply to their decisions. Additionally, inducing time consequences instead of money consequences leads to unobservable variations in the opportunity cost of time. On the other hand, strengthening internal validity in field experiments also comes at a weakening of external validity. Controlling incentives and contextual parameters can be seen as very artificial by the participants. Experiments conducted in virtual reality simulations, like in driving simulators, lie someplace in between the pure lab and pure field experiment. Though natural motivations (and consequences) do not exist, many of the natural cues from the field can be implemented in the simulator, in a natural-like fashion, but stricter experimental protocols that lead to a higher level of internal validity than in the field can be applied.

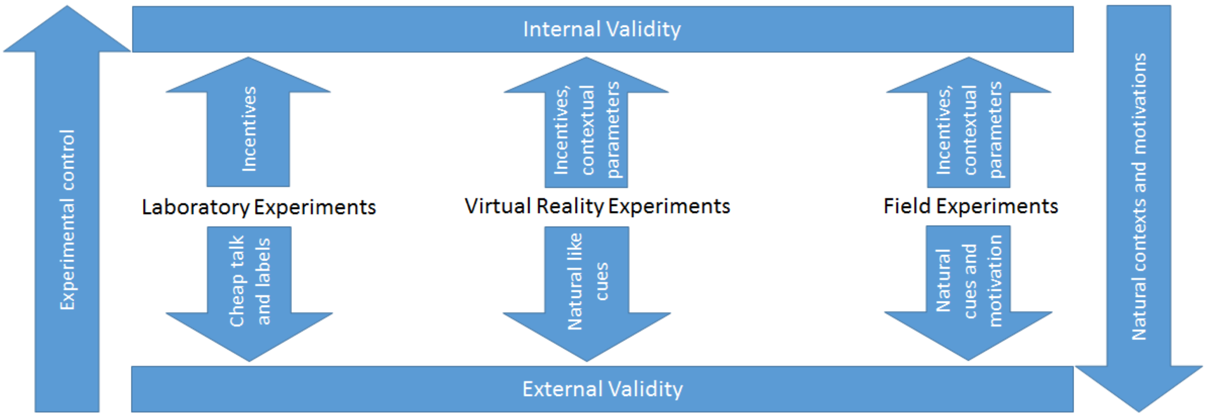

FIGURE 3: Types of Experiments and relation to internal and external validity

Based on the lab experiments reviewed here we identify five main strategies that have been employed to increase external validity. These include:

1) Labels with Field Referents
2) Virtual reality, driving simulators and realistic cost functions
3) Inducing time instead of money in the incentive structure
4) The use of field participants from relevant populations

The first three are discussed in the previous section. The use of field participants needs to be discussed since most laboratory experiments have relied on students as participants. The argument made is that when the purpose of the experiment is theory testing, students provide a relevant and low cost participant pool since “students are people too”. When it comes to theory testing in experiments with a high internal validity, students appear to behave in ways very similar to non-students (e.g., Huck and Mueller 2012 or Bosch-Domenech et al. 2002). When external validity is increased through any of the first three strategies listed above, or in other ways, it is likely that field participants are affected differently than students because of their greater familiarity and experiences with the contexts referred to. Additionally, if the purpose of the experiment is to elicit preference parameters, such as attitudes to risk or time delays, students may not match the preferences of the target population in the field, necessitating the use of field participants.

However, some questions can only be studied in the field. One especially important example is that of evaluating value of time. The value of time reflects the opportunity cost of the time use, and a participant may have a very different value of time for work related commutes during rush hours than she has on a Saturday when she is taking a more leisurely drive for recreational purposes. Thus, if the interest in value of time is derived from the wish to establish optimal congestion pricing during rush hours, then it should be inferred from driving choices on the routes and times that correspond to rush hour traffic.

While beliefs about context specific travel times can be elicited in a laboratory, using relevant field populations with familiarity of the routes, it is also possible to study the subjective process by which experience influences beliefs in the lab. People may differ both in how strongly they hold onto prior beliefs, how sensitive they are to experiences or to information, and how quickly they forget older experiences, and these factors may be specific to the individual participant but relatively independent of the context. These factors are therefore amenable to study in a lab environment. In a similar way, attitudes to travel time unreliability may not vary across context but may be intrinsic to the participant. If so, an understanding of the distribution of such attitudes can be based on data collected in lab experiments, but with relevant field populations. Further, certain cognitive characteristics of individuals may be context independent, and therefore amenable to be studied in lab settings.

Similarly, use of virtual reality simulations and driving simulators can provide a bridge between the internal validity of the lab experiment and the external validity of the field experiment. Many of the factors that generate internal validity can be maintained, such as dominance and saliency, while many of the field cues and experiences that may trigger specific decision heuristics can be introduced.

4.5 Scientific Requirements of Experiments

Experimental economists have laid out scientific requirements to ensure replicability of data, reliable participant pools, and incentive compatible observation methods:

- No data should be collected through deception (Hertwig and Ortmann 2001; Ortmann and Hertwig 2002; Hertwig and Ortmann 2008). This is critical since many experiments deal with studying choice behavior so that participants maximize expected utility or non-expected utility frameworks. Experimental environments where participants are frequently deceived destroys the saliency of the incentives and instructions offered. Non-deception is an ethical principle aimed at creating a public good to the scientific community where participants are not contaminated by previous deception.
- All instructions and data generated should be readily available to researchers so that results can be replicated and confounding effects of experimental design explored further.

- Ethical human participant principles should be adhered to such as: non-exploitation of vulnerable populations, no deception of participants with respect to potential harmful effects of participation, and no coercion.

5. CONCLUSIONS

The adherence to Induced Value Theory has made it possible for economics experiments to achieve a high degree of internal validity to study causality in laboratory and field settings. However, experimenters have always strived to achieve external validity. New innovations in virtual reality and driving simulation have made these technologies cheaper and accessible for laboratory experiments. The contextual realism and immersion provide subjects a high degree of presence in reality. Furthermore, smart phones, smart cards and connected technology are providing access to higher resolution data and ability to conduct more controlled field experiments more cheaply than ever before. This has made possible the ability to conduct a continuum of such experiments.

A good experimental design will weigh the requirements for internal and external validity from the perspective of the research question. While it can be tempting to desire a field experiment to address specific policy questions, it is important to understand that many field aspects can be modelled in the lab at a much lower cost and with higher internal validity. The internal validity also allows researchers to more fundamentally understand the causality behind both effects on mean shifts in behaviour and on changes in the variance in aggregate behaviour.

Each method used to study transportation behaviour has its strength and weaknesses. While experimental economics can achieve high internal validity, it comes at a cost due to the induced monetary values. Stated choice methods have the advantage that tasks can easily be linked to familiar settings with presentations of proposed changes linked to the present state of the transportation system. By applying both methods to generate evidence on behavioural responses to policy changes, we can collect more observations and reduce biases and noise in the data, enabling a more solid understanding of the phenomena.

Apart from addressing questions of ongoing interest, such as effects of congestion pricing on route-, mode- and departure time choices, there are a number of new questions in transportation that will need to be addressed and that can benefit from the internal validity of induced value experiments. For instance, exploring methods to enhance cooperation in transportation systems (Klein and Ben-Elia, 2016), as well as the role of subjective beliefs and risk attitudes on adoption of new transportation and mobility options (eg. Autonomous vehicles). Using incentives from methods in experimental economics, and contextual realism through laboratory, field and virtual reality experiments would help provide experimenters tools to gain behavioural insights to inform policies.

5.1 Experiments and Technology

Technology can be used to improve on the data collected through experiments, and experiments can also be used to evaluate the behavioural effects of new technology. GPS technology has implied a decreased reliance on self-reporting of driving choices in the field, allowing instead the experimenter to directly observe the choices by participants. On-board computer data allows the experiment to observe driving behaviour even more accurately. However, with an increased interest of insurance companies and regulatory agencies in accessing on-board computer data, experiments can also be used to evaluate these uses of technology. Interesting research questions include the acceptability of being monitored so directly in different social and cultural settings. As fully automated vehicle technologies become more and more feasible in everyday traffic, experiments can be used to track behaviour of drivers in non-automated vehicles as they interact with the automated ones, or as human drivers choose to manually override automation. The use of driving simulators can allow experimenters to observe such behaviour in safe environments at zero accident costs. Driving simulator technology that allows simultaneous and interactive choices by multiple simulator drivers is now available, opening up new research possibilities. As simulator technology improves further, perhaps large-scale simulations of traffic systems will allow incorporation of individual human-operated vehicles to test behavioural assumptions underlying simulated vehicle traffic. Mobile phone technology can be used to provide real-time information to drivers, and behavioural evaluations of such technologies can be conducted via controlled experiments.

This review points to important experimental design elements that need attention as new technologies get incorporated in experiments. We believe experimental economics is an important research tool, complementary to other data collection methods, and expect to see a great deal of interesting future research.

Acknowledgements

We acknowledge that this work was supported by the Federal Highway Administration under Agreement No. DTFH61-09-H-00012. We are grateful to late Prof. Matthew Karlaftis who encouraged us to work on carrying out a comprehensive review of this subject. We also thank the editor and the anonymous reviewers for their comments that has helped improve this manuscript.

REFERENCES

1. Abdellaoui, M., Kemel, E., 2014. Eliciting Prospect Theory When Consequences Are Measured in Time Units: "Time Is Not Money". Management Science, 60(7), 1844-59.
2. Agerholm, N., Waagepetersen, R., Tradisauskas, N., Harms, L., Lahrmann, H. I., 2008. "Preliminary results from the Danish Intelligent Speed Adaptation project Pay as You Speed." IET Intelligent Transport Systems, 2(2), 143-153.
3. Alonso, W., 1964. Location and land use. Cambridge, Harvard University Press.
4. Anas, A., Arnott, R., Small, K. A., 1998. Urban Spatial Structure. Journal of Economic Literature, 36 (3), 1426-1464.
5. Andersen, S., Dixit, V. V., Martinez-Correa, J., Rutström, E. E., 2014. When the Field is not Enough. Using Complementary Field and Lab Experiments to Investigate Congestion Pricing Responses. DBEL-WP 1402, Georgia State University, March.
6. Andersen, S., Cox, J. C., Harrison, G. W., Lau, M. I., Rutström, E. E., Sadiraj, V., 2012. Asset Integration and Attitudes to Risk: Theory and Evidence. Working Paper. Available at: http://excen.gsu.edu/workingpapers/GSU_EXCEN_WP_2012-12.pdf
7. Andersen, S., Dixit, V. V., Harrison, G. W., Harb, R., Lau, M. I., Radwan, E., Rutström, E. E., Tarr, R., 2011. Experiments on Driving Under Uncertain Congestion Conditions and the Effects on Traffic Networks from Congestion Initiatives. Report to U.S. Federal Highway Administration. DBEL-WP 1101, Georgia State University, November.
8. Andersen, S., Harrison, G. W., Lau, M. I., Rutström, E. E. , 2010, Preference Heterogeneity in Experiments: Comparing the Lab and the Field, Journal of Economic Behavior and Organization, 73, 2010, 209-224.
9. Anderson, L., Holt, C. A., Reiley, D., 2008. Congestion Pricing and Welfare: An entry experiment. In Cherry, Koll, Shogren (eds), Environmental Economics, Experimental Methods, chapter 15, 280-92.
10. Arnott, R., De-Palma, A., Lindsey, R., 1993. A Structural Model of Peak-Period Congestion: A Traffic Bottleneck with Elastic Demand. American Economic Review, 83(1), 161-179.
11. Arnott, R., Small, K., 1994. The Economics of Traffic Congestion. American Scientist, 82(1), 446-455.
12. Avineri, E., Prashker, J. N., 2006. The impact of travel time information on travellers' learning under uncertainty. Transportation, 33(4), 393–408.
13. Aziz, H. M., Romero, J., Ukkusuri, S. V., 2015. Understanding short-term travel behavior under personal mobility credit allowance scheme using experimental economics. Transportation Research Part D, 36, 121-137.
14. Ben-Elia, E., Avineri, E., 2015. Response to Travel Information: A Behavioural Review. Transport Reviews, 35(3), 352-377.
15. Ben-Elia, E., Erev, I., Shiftan, Y., 2008. The combined effect of information and experience on drivers' route-choice behaviour. Transportation, 35(2), 165–177.
16. Ben-Elia, E., Ishaq, R., Shiftan, Y., 2013. If only I had taken the other road…: Regret, Risk and Reinforced Learning in Informed Route-Choice. Transportation, 40(2), 269-293.

17. Ben-Elia, E., Shiftan, Y., 2010. Which road do I take? A learning-based model of route-choice behaviour with real-time information. Transportation Research Part A, 44(4), 249-264.
18. Ben-Elia, E., Ettema, D., 2011a. Rewarding rush-hour avoidance: A study of commuter's travel behavior. Transportation Research Part A, 45(7), 567-582.
19. Ben-Elia, E., Ettema, D., 2011b. Changing commuters' behavior using rewards: A study of rush-hour avoidance. Transportation Research Part F, 14(5), 354-368.
20. Bergantino, A. S., De Carlo, A., Morone, A., 2015. Individuals' behaviour with respect to parking alternatives: a laboratory experiment. MPRA Paper 63815, University Library of Munich, Germany.
21. Bertrand, M., Djankov, S., Hanna, R., Mullainathan, S., 2007. Obtaining a driver's license in India: an experimental approach to studying corruption. The Quarterly Journal of Economics, 122(4), 1639-1676.
22. Binmore, K., Shaked, A. 2010. Experimental economics: Where next? Journal of Economic Behavior and Organization, 73(1), 87-100.
23. Bliemer, M., Denant-Boemont, L., Hammiche, S., Hensher, D., Mulley C., 2013. Transport Costs and the Structure of the City: A Laboratory Experiment. 13th World Conference on Transport Research WCTR 2013, Rio de Janeiro, Brazil, 18th July 2013.
24. Blumenschein, K., Blomquist, G., Johannesson, M., Horn, N., Freeman, P., 2008. Eliciting willingness to pay without bias: evidence from a field experiment. Econ. J. 118(525), 114–137.
25. Bohm, P., 2002. Pitfalls in experimental economics. In F. Andersson and H. Holm (Eds.), Experimental economics: financial markets, auctions, and decision making. Dordrecht: Kluwer Academic
26. Bohm, P., 1984. Are there practicable demand-revealing mechanisms? In H. Hanusch (Ed.), Public finance and the quest for efficiency. Detroit: Wayne State University Press.
27. Bolderdijk, J. W., Knockaert, J., Steg, E. M., Verhoef, E. T., 2011. Effects of Pay-As-You-Drive vehicle insurance on young drivers speed choice: Results of a Dutch field experiment. Accident Analysis and Prevention, 43(3), 1181-86.
28. Bosch-Domenech, A., Montalvo, J. G., Nagel, R., Satorra, A., 2002. One, two, (three), infinit...: Newspaper and lab beauty-contest experiments. American Economic Review, 92(5), 1687-1701.
29. Brilon, W., Geistefeldt, J., Regler, M., 2005. Reliability of Freeway Traffic Flow: A stochastic Concept of Capacity. Proceedings of the 16th International Symposium on Transportation and Traffic Theory, 125-144.
30. Brown-Kruse, J., Schenk, D. J., 2000. Location, Cooperation, and Communication: An Experimental Examination. International Journal of Industrial Organization, 18(1), 59-80.
31. Bureau of Public Roads (BPR), 1964. Traffic Assignment Manual. U.S. Dept. of Commerce, Urban Planning Division, Washington D.C.
32. Camerer, C., Lovallo, D., 1999. Overconfidence and excess entry: An experimental approach. American economic review, 89(1), 306-318.
33. Cavailhes, J., Gaigné, C., Tabuchi T., Thisse, J. F., 2007. Trade and the structure of cities. Journal of Urban Economics, 62(3), 383-404.

34. Champ, P., Moore, R., Bishop, R., 2009. A comparison of approaches to mitigate hypothetical bias. Agric. Resour. Econ. Rev. 38(2), 166-180.
35. Chmura, T., Pitz, T., 2004. Minority Game - Experiments and Simulations of Traffic Scenarios." Bonn Econ Discussion Papers bgse23_2004, University of Bonn, Germany.
36. Chmura, T., Pitz, T., 2004. An Extended Reinforcement Algorithm for Estimation of Human Behaviour in Congestion Games. Bonn Econ Discussion Papers bgse24_2004, University of Bonn, Germany.
37. Chmura, T., Pitz, T., 2007. An Extended Reinforcement Algorithm for Estimation of Human Behaviour in Congestion Games. Journal of Artificial Societies and Social Simulation, 10 (2), http://jasss.soc.surrey.ac.uk/10/2/1.html, Paper 24/2004, Bonn Graduate School of Economics, December 2004.
38. Chmura, T., Pitz, T., 2006. Successful Strategies in Repeated Minority Games. Physica A, 363, 477-480.
39. Cox, J. C., Robertson, B., Smith, V. L., 1982. Theory and Behaviour of Single Object Auctions. In V. L. Smith, ed., Research in Experimental Economics, Greenwich: JAI press, 1982.
40. Cox, J. C., Smith, V. L., Walker, J. M., 1983. Test of Heterogeneous Bidder's Theory of First Price Auctions. Economic Letters, 12(3-4), 207-212.
41. Cox, J. C., Smith, V. L., Walker, J. M., 1985. Experimental Development of Sealed-Bid Auction Theory: Calibrating Controls for Risk Aversion. American Economic Review, 75(2), 160-65.
42. Cox, J. C., Smith, V. L., Walker, J. M., 1988. Theory and Individual Behaviour of First Price Auctions. Journal of Risk and Uncertainty. 1(1), 61-99.
43. Cox, J. C., Offerman, T., Olson, M. A., Schram, A. J. H. C., 2003. Competition For Versus On The Rails: A Laboratory Experiment. International Economic Review 43(3), 709-736.
44. Cox, J. C., Sadiraj, V., Schmidt, U., 2015. Paradoxes and Mechanism for Choice under Risk. Experimental Economics, 18(2), 215-250
45. Cummings, R. G.; G. Harrison; and E. E. Rutström (1995). "Homegrown Values and Hypothetical Surveys: Is the Dichotomous Choice Approach Incentive Compatible?" American Economic Review, 85(1), March 1995, 260-266.
46. Daganzo, C. F., Sheffi, Y. 1977. On stochastic models of traffic assignment. Transportation Science 11 (3), 253–274.
47. Daniel, T. E., Gisches, E. J., Rapoport, A., 2009. Departure Times in Y-Shaped Traffic Networks with Multiple Bottlenecks. The American Economic Review, 99 (5), 2149-2176
48. DeAngelo, G., Charness, G., 2012. Deterrence, expected cost, uncertainty and voting: Experimental evidence. Journal of Risk and Uncertainty, Springer, 44(1), 73-100.
49. Dechenaux, E., Mago, S. D., Razzolini, L., 2014. Traffic congestion: an experimental study of the Downs-Thomson paradox. Experimental Economics, 17(3), 461-487.
50. Denant-Boemont, L., Hammiche, S. 2010. Flexibility of Individual Decisions in a Real-Option Setting: An Experimental Study of Transport Choice. Journal of Intelligent Transport Systems, 14(3), 1-14.

51. Denant-Boemont, L., Hammiche, S., 2012. Downs-Thomson Paradox and Public Transit Capacity in the Laboratory. ITLS WP-12-10.
52. Denant-Boemont, L., 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.
53. Denat-Boemont, L., Petiot, R., 2003. Information Value and Sequential Decision-Making in a Transport Setting: An Experimental Study. Transportation Research Part B, 37, 365-386.
54. Devetag, G., Ortmann, A., 2007. When and why? A critical survey on coordination failure in the laboratory. Experimental Economics, 10 (3), 331-344.
55. Dixit, V. V., Rutström, E., Mard, S., Zielske, R., 2010. Factors Influencing Outcomes of Referendums for Transit Funding. Transportation Research Record, 2143, 44-47.
56. Dixit, V.V., Harb, R., Martinez-Correa J., Rutström, E. E., 2015. Measuring Risk Aversion to Guide Transportation Policy: Contexts, Incentives, and Respondents, Transportation Research Part A Policy and Practice, vol 80, October, pp. 15-34.
57. Dixit, V. V., Harrison, G., Rutström, E. E., 2014. Estimating the Subjective Risks of Driving Simulator Accidents. Accident Analysis and Prevention, 62, 63–78.
58. Dixit, V. V., Denant-Boemont , L., 2014. Is equilibrium in transport pure Nash, mixed or Stochastic? Transportation Research Part C: Emerging Technologies, 48, 301-310.
59. Dufwenberg, M., Harrison, G. W., 2008. Peter Bohm – Father of Field Experiments. Experimental Economics 11(3), 213-20.
60. Eliasson, J., Hultkrantz, L., Nerhagen, L., Rosqvist, L. S., 2009. The Stockholm congestion – charging trial 2006: Overview of effects. Transportation Research Part A: Policy and Practice, 43(3), 240–250
61. Fifer, S., Rose, J., Greaves, S., 2014. Hypothetical bias in Stated Choice Experiments: Is it a problem? And if so, how do we deal with it? Transportation Research Part A: Policy and Practice, 61, 164-177
62. Fiore, S. M., Harrison, G. W., Hughes, C. E., Rutström, E. E., 2009. Virtual Experiments and Environmental Policy. Journal of Environmental and Economic Management, 57(1), 65-86.
63. Gabuthy, Y., Neveu, M., Denant- Boèmont, L., 2006. The coordination problem in a structural model of peak-period congestion: An experimental study. Review of Network Economics, 5(2), 273-298.
64. Goeree, J. K., Holt, C. A., 2005. An Explanation of Anomalous Behavior in Models of Political Participation. American Political Science Review, 99(2), 201-213.
65. Goeree, J. K., Holt, C. A., Palfrey, T., 2008. Quantal Response Equilibrium. New Palgrave Dictionary of Economics, Second Edition, Editors Steven N. Durlauf and Lawrence E. Blume, Palgrave Macmillan, 2008.
66. Gomes, N., Merugu, D., O'Brien, G., Mandyam, C., Yue, T., Atikoglu, B., Albert, A., Fukumoto, N., Liu, H., Prabhakar, B., Wischik, D., 2012. Steptacular: an Incentive Mechanism for Promoting Wellness. NetHealth, Comsnets Workshop on Networked Healthcare Technology, January 2012.
67. Greenbaum, T., 2000. Moderating Focus Groups. Thousand Oaks, California: Sage Publications, Inc. ISBN 0-7619-2044-7

68. Hartman, J. 2012. A Route Choice Experiment with an Efficient Toll. Networks and Spatial Economics, 12(2), 205-222.
69. Harrison, G. W. 1989. Theory and Misbehavior of First-Price Auctions. American Economic Review, 79(4), 749-762.
70. Harrison, G. W., Rutström, E. E., 2001. Doing it both ways – experimental practice and heuristic context. Behavioral and Brain Sciences, 24(3), 413-414.
71. Harrison, G. W., Rutström, E. E., 2008. Risk Aversion in the Laboratory, in James C. Cox, Glenn W. Harrison (ed.) Risk Aversion in Experiments (Research in Experimental Economics, Vol. 12) Emerald Group Publishing Ltd., 41-196..
72. Harrison, G. W., Swarthout, T., 2014. Experimental payment protocols and the Bipolar Behaviorist. Theory and Decision, 77 (3), 423-438.
73. Hensher, D., 2010. Hypothetical bias, choice experiments and willingness to pay. Transportation Research Part B: Methodological, 44(6), 735-752.
74. Hertwig, R., Ortmann, A., 2001. Experimental Practices in Economics: A Methodological Challenge for Psychologists? Behavioral and Brain Sciences, 24(3), 383–451.
75. Hertwig, R., Ortmann, A., 2008. Deception in Experiments: Revisiting the Arguments in Its Defense. Ethics and Behavior, 18(1), 59 - 92.
76. Holguín-Veras, J., Thorson, E., 2003. The role of experimental economics in freight transportation research: preliminary results of experimentation. Paper presented at the 2003 European Transport Conference, Strasbourg.
77. Holguín-Veras, J., Silas, M., Polimeni, J., Cruz, B., 2007. An Investigation on the Effectiveness of Joint Receiver-Carrier Policies to Increase Truck Traffic in the Off-Peak Hours. Networks and Spatial Economics, 7(3), pp 277-295
78. Holt, C. A., Laury, K. S., 2002. Risk Aversion and Incentive Effects. American Economic Review, 92(5), 1644-1655.
79. Hotelling, H., 1929. Stability in Competition. Economic Journal 39 (153), 41–57
80. Huck, S., Müller, W., Vriend, N. J., 2002. The east end, the west end, and King’s Cross: on clustering in the four-player hotelling game. Economic Inquiry, 40(2), 231-240.
81. Huck, S., Normann, H., Oechssler, J., 2004. Two are few and four are many: number effects in experimental oligopolies. Journal of Economic Behavior & Organization, 53 (4), 435-446.
82. Huck, S. and W. Müller, W., 2012. Allais for all: Revisiting the paradox in a large representative sample. Journal of Risk and Uncertainty, 44(3), 261-293.
83. Hultkrantz, L., Lindberg, G., 2011. “Accident Cost, Speed and Vehicle Mass Externalities, and Insurance.” International Transport Forum Discussion Papers 2011/26, OECD Publishing.
84. Iida, Y., Akiyama, T., Uchida, T., 1992. Experimental analysis of dynamic route choice behaviour. Trans. Res. B, 26, 17-32.
85. Innocenti, A., Lattarulo, P., Pazienza, M. G., 2013. Car stickiness: Heuristics and biases in travel choice. Transport Policy, 25, 158–168.
86. Isacsson, G., Nilsson, J. E., 2003. An experimental comparison of track allocation mechanisms in the railway industry. Journal of Transport Economics and Policy, 37(3), 357-382.

87. Janusch, N., Kroll, S., Goemans, C., 2015. Acceptability of Congestion Pricing: Worldviews Versus Experience. Canadian Economics Association 2015 Annual Conference; Toronto, Ontario, May 30, 2015.
88. Johannesson, M., 1999. Calibrating hypothetical willingness to pay responses. J. Risk Uncertainty, 18(1), 21–32.
89. Kemel, E., Paraschiv, C., 2013. Prospect theory for joint time and money consequences in risk and ambiguity. Transportation Research Part B, 56, 81–95
90. Klein, I., Ben-Elia, E., 2016. Emergence of cooperation in congested road networks using ICT and future emerging technologies: A game-based review. Transportation Research Part C: Emerging Technologies, 72, 10-28.
91. Lammers, F., Schiller, J., 2010. Contract design and insurance fraud: An experimental investigation. FZID Discussion Papers 19-2010, University of Hohenheim, Centre for Research on Innovation and Services (FZID).
92. Ledyard, J. O., Olson, M., Porter, D., Swanson, J. A., Torma, D. P., 2002. The first use of a combined-value auction for transport services. Experimental Economics in Practice, 31(5), 4-12.
93. Li, C. Z., Mattsson, L., 1995. Discrete choice under preference uncertainty: an improved structural model for contingent valuation. J. Environ. Econ. Manage. 28(2), 256–269.
94. List, J., Lucking-Reiley, D., 2000. Demand Reduction in Multiunit Auctions: Evidence from a Sportscard Field Experiment. American Economic Review, 90(4), 961-972
95. Liu, F., 2013. Essays on Accident Forgiveness in Automobile Insurance. Ph. D. Dissertation, Georgia State University.
96. Loewenstein, G., 1999. Experimental Economics from the Vantage-Point of Behavioural Economics. Economic Journal, 109(453), F25-F34.
97. Loomes, G., 1999. Experimental Economics: Introduction. Economic Journal, 109(453), FI-F4.
98. Lunander, A., Nilsson, J. E., 2004. Experimental Tests of Alternative Mechanisms to Procure Multiple Contracts. Journal of Regulatory Economics, 25(1), 39-58.
99. Mak, V., Gisches, E. J., Rapoport, A., 2015. Route vs. Segment: An Experiment on Real-Time Travel Information in Congestible Networks. Production and Operations Management, 24(6), 947–960.
100. Mahmassani, H. S., Jou, R. C., 2000. Transferring Insights into Commuter Behavior Dynamics from Laboratory Experiments to Field Surveys. Transportation Research part A, 34, 243–260.
101. Mahmassani, H. S., Herman, R., 1990. Interactive Experiments for the Study of Tripmaker Behaviour Dynamics in Congested Commuting Systems. Developments in Dynamic and Activity-Based Approaches to Travel Analysis, P. Jones (ed.), 272–298, Gower, Aldershot, UK.
102. Mahmassani, H. S., Srinivasan, K. K., 2004. Experiments with Route and Departure Time Choices of Commuters Under Real-Time Information: Heuristics and Adjustment Processes. Human Behavior and Traffic Networks, M. Schreckenburg and R. Selten (eds.), 97–132, Springer-Verlag, Berlin, Germany.
103. McKelvey, R., Palfrey, T. 1995. Quantal Response Equilibria for Normal Form Games. Games and Economic Behavior. 10, 6–38.

104. McKelvey, R., Palfrey, T. 1998. Quantal Response Equilibria for Extensive Form Games. Experimental Economics, 1(1), 9-41
105. Meinhold, G., Pickhardt, M., 2009. Road Pricing and Excess Burden Avoidance: Evidence from an Experiment with Braess Paradox, working paper.
106. Merugu, D., Prabhakar, B., Rama, N., 2009. An Incentive Mechanism for Decongesting the Roads: A Pilot Program in Bangalore. NetEcon, ACM Workshop on the Economics of Networked Systems, July 2009.
107. Ming, T., 2015. Estimating Subjective Beliefs in Naturalistic Tasks with Limited Information. 56th Annual Transportation Research Forum, Atlanta, March 12-14, 2015.
108. Moore, R., Bishop, R.C., Provencher, B., Champ, P.A., 2010. Accounting for respondent uncertainty to improve willingness-to-pay estimates. Can. J. Agric. Econ. Rev. can. d'agroecon. 58(3), 381–401.
109. Morgan, J., Orzen, H., Sefton, M., 2009. Network architecture and traffic flows: Experiments on the Pigou-Knight-Downs and Braess Paradoxes. Games and Economic Behaviour, 66(1), 348-372.
110. Morrison, M., Brown, T., 2009. Testing the effectiveness of certainty scales, cheap talk, and dissonance-minimization in reducing hypothetical bias in contingent valuation studies. Environ. Resour. Econ. 44(3), 307–326.
111. Nielsen, O. A., Jovicic, G., 2003. The AKTA Road Pricing Experiment in Copenhagen. paper presented at the 10th International Conference on Travel Behaviour Research, Lucerne, August 2003.
112. Ortmann, A., 2016. Episodes from the Early History of Experimentation in Economics. in A. Svorencik & H. Maas (eds.), pp.195 – 217, Witness Seminar on the Emergence of a Field, Cham, Switzerland: Springer 2016.
113. Ortmann, A., Hertwig, R., 2002. The Costs of Deception: Evidence from Psychology. Experimental Economics, 5:111–131.
114. Pluntke, C., Prabhakar, B. 2013. INSINC: A Platform for Managing Peak Demand in Public Transit. JOURNEYS, Land Transport Authority Academy of Singapore, September 2013.
115. Ramadurai, G., Ukkusuri, S. V., 2007. Dynamic Traffic Equilibrium: Theoretical and Experimental Network Game Results in the Single Bottleneck Model. Transportation Research Record: Journal of the Transportation Research Board, 2029, 1-13.
116. Rapoport, A., Mak, V., Zwick, R., 2006. Navigating congested networks with variable demand: experimental evidence. Journal of Economic Psychology, 27(5), 648-666.
117. Rapoport, A., Kugler, T., Dugar, S., Gisches, E. J., 2009. Choice of routes in congested traffic networks: Experimental tests of the Braess Paradox. Games and Economic Behavior, 65(2), 538-571.
118. Rapoport, A., Gisches, E., Daniel, T., Lindsey, R., 2014. Pre-trip information and route-choice decisions with stochastic travel conditions: Experiment. Transportation Research Part B, 68, 154-172.
119. Rosenthal, R. W., 1973. A class of games possessing pure-strategy Nash equilibria. International Journal of Game Theory, 2(1), 65–67.

120. Rydval, O., Ortmann, A., 2004. How financial incentives and cognitive abilities affect task performance in laboratory settings: an illustration. Economics Letters, 85(3), 315-320.
121. Sanchez-Vives, M., M. Slater, 2005. From presence to consciousness through virtual reality. Nature Reviews Neuroscience, 6, 332-339.
122. Schneider, K., Weimann, J., 2004. Against all odds: Nash equilibria in road pricing experiment. In M. Schreckenberg and R. Selten (Eds.), Human Behaviour and Traffic Networks. Berlin: Springer, 133-153.
123. Small, K. A., Winston, C., Yan, J., 2005. Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability. Econometrica, 73(4), 1367-1382
124. Smith, V. L. 1976. Experimental Economics: Induced Value Theory. American Economic Review, 66(2), 274–279.
125. Smith, V. L. 1982. Microeconomic Systems as an Experimental Science. The American Economic Review, 72(5), 923-955.
126. Spicer, M., Becker, L. A., 1980. Fiscal Inequity and Tax Evasion: An Experimental Approach. National Tax Journal, 33(2), 171-175.
127. Srinivasan, K. K., Mahmassani, H. S., 1999. Role of Congestion and Information in Trip-Makers' Dynamic Decision Processes: Experimental Investigation. Transportation Research Record, 1676, 44–52.
128. Tørnblad, S.H., Kallbekken, S., Korneliussen, K., Mideksa, T. K., 2014. Using mobility management to reduce private car use: Results from a natural field experiment in Norway. Transport Policy, 32, 9-15.
129. Tsang, M. 2015. Estimating Subjective Beliefs in Naturalistic Tasks with Limited Information. Economic Science Association North-American Meeting, Dallas Texas. Oct 22-24, 2015
130. Vickrey, W., 1969. Congestion Theory and Transport Investment. American Economic Review, 59(2), 251-260.
131. Wardrop, J., 1952. Some theoretical aspects of road traffic research. Proceedings of the Institution of Civil Engineers, Part II(1), 325-378.
132. Yan, X., M. Abdel-Aty, E. Radwan, X. Wang, P. Chilakapati, 2008. Validating a driving simulator using surrogate safety measures. Accident Analysis & Prevention, 40(1), 274-288.
133. Ziegelmeyer, A., Koessler, F., My, K. B., Denant-Boemont, L., 2008. Road Traffic Congestion and Public Information. Journal of Transport Economics and Policy, 42(1), 43-82.