



## **WORKING PAPER**

**ITLS-WP-20-08**

### **Route Choice Behaviour: Stated Choices and Simulated Experiences**

**By**

**Muhammad Fayyaz<sup>a</sup>, Michiel Bliemer<sup>a</sup>, Matthew Beck<sup>a</sup>, Stephane Hess<sup>b</sup> and Hans van Lint<sup>c</sup>**

<sup>a</sup> Institute of Transport and Logistic Studies (ITLS), The University of Sydney Business School, Sydney Australia

<sup>b</sup> Choice Modelling Centre, Institute for Transport Studies, University of Leeds, United Kingdom

<sup>c</sup> Department of Transport and Planning, Delft University of Technology, the Netherlands

**April 2020**

**ISSN 1832-570X**

## **INSTITUTE of TRANSPORT and LOGISTICS STUDIES**

The Australian Key Centre in  
Transport and Logistics Management

The University of Sydney

*Established under the Australian Research Council's Key Centre Program.*



**NUMBER:** Working Paper ITLS-WP-20-08

**TITLE:** **Route Choice Behaviour: Stated Choices and Simulated Experiences**

**ABSTRACT:** Surveys with stated choice experiments (SCE) are widely used to examine route choice behaviour in hypothetical choice contexts and to derive values of time and reliability for transport project appraisal purposes. In contrast to revealed choices, stated choices do not let participants experience (the consequences of) any attribute, which is one of the reasons why the external validity of SCE outcomes is often questioned. In this paper, we investigate the impact of simulated experiences on behaviour in a route choice context. We recruited 74 people who completed both a typical SCE and an incentive compatible driving simulator experiment (DSE), where the latter required respondents to experience the travel time of their chosen route and actually pay any toll costs associated with the choice of a tolled road. The choices are analysed via a heteroscedastic latent class model. Compared to the SCE, in the DSE, participants selected the tolled road less often, suggesting that having to pay actual money changes stated preferences. Furthermore, we found large variations in sensitivity to toll cost across participants. On the other hand, we found only minor differences in preferences towards travel time and travel time unreliability between SCE and DSE.

**KEY WORDS:** *Hypothetical route choice contexts; Route choice data collection techniques; Stated choice experiments; Incentive compatible driving simulator experiments; Route choice behaviour; Value of travel time; Value of travel time reliability.*

**AUTHORS:** **Fayyaz, Bliemer, Beck, Hess and Lint**

**Acknowledgements:** This research is funded by Australian Research Council grants DP150103299 and DP180103718. Stephane Hess also acknowledges support for his inputs by the European Research Council through the consolidator grant 615596-DECISIONS.

**CONTACT:**

INSTITUTE OF TRANSPORT AND LOGISTICS STUDIES  
(H73)

The Australian Key Centre in Transport and Logistics  
Management

The University of Sydney NSW 2006 Australia

Telephone: +612 9114 1824

E-mail: [business.itlsinfo@sydney.edu.au](mailto:business.itlsinfo@sydney.edu.au)

Internet: <http://sydney.edu.au/business/itls>

**DATE:**

April 2020



## 1. Introduction

Route choice has been a topic of study for many decades in order to better understand the importance of various route attributes and to forecast behaviour in networks for transport planning and management purposes. Stated preference (SP) and revealed preference (RP) are the two common data types used to examine travellers' route choice behaviour. SP data typically are collected with a stated choice experiment (SCE), in which respondents are asked to make choices in a series of hypothetical choice tasks. In contrast, RP data consists of route choices observed in the field, for example where drivers are tracked using mobile phones or other Global Positioning System (GPS) devices, remote sensing, or using driver-reported route information in interviews and questionnaires. Using any of these types of data, it is possible to formulate models and estimate key factors influencing behaviour, and particularly sensitivities to journey characteristics that allow an analyst to determine for example the value of travel time (VTT) and the value of travel time reliability (VOR). VTT is the monetary value drivers assign to travel time changes, and VOR is the monetary value assigned to a change in travel time variability (unreliability). For the past five decades, the VTT has been considered an important value in transport policies and transport projects appraisal (Abrantes & Wardman, 2011). VTT serves two purposes, namely (i) as an input variable in cost-benefit analysis (CBA) of transport infrastructure projects, and (ii) as an explanatory variable in transport forecasting models (Shires & de Jong, 2009). More recently (for the last two decades), VOR has also received considerable attention in the CBA of transport projects and policies (de Jong & Bliemer, 2015).

VTT and VOR are often estimated using SP rather than RP data<sup>1</sup>. For example, in the UK, a large number of SCEs were designed for estimating VTT and VOR for different transport modes and trip purposes (Hess et al., 2017). Although it is assumed that behaviour captured by SCEs reflects real-world behaviour, the hypothetical nature of SCEs may lead to biased results (Beck et al., 2016; Fifer et al., 2014). One of the potential reasons for hypothetical bias in typical SCEs is that "participants may not experience strong incentives to expend the cognitive efforts needed to provide researchers with an accurate answer" (Ding et al., 2005, p. 68). A SCE would therefore be incentive compatible only if it provides an incentive for participants to nudge them to truthfully reveal their preference towards an attribute. Experiments designed specifically to be incentive compatible are widely conducted in experimental economics and are often referred to as an economic experiment. Conversely, the tendency to simplify SCE settings, a practice that prevails despite extensive criticism (cf. Hess et al., 2020a), may lead to respondents enhancing the information in an unobserved manner to make up for the lack of experienced stimuli.

Economic experiments differ from SCEs (and experiments in other fields) in the sense that the incentives developed are explicitly designed such that they will reveal the participants' true preferences. To design such incentive structures and implement behavioural control over participants, economic experiments rely on induced value theory (Smith, 1976). According to Smith, three conditions must be satisfied to explore incentive compatible behaviour: monotonicity, dominance and salience. Monotonicity requires that participants prefer more reward over less reward and this condition is satisfied if money is used as a reward medium. Dominance means that the participants' utility in the experiment should mainly come from the incentives while other factors are of negligible influence. Salience implies that incentives should depend on participants' actions. For incentive compatibility, salience is the most relevant condition and can be satisfied when participants' monetary reward depends directly on choices made during the experiment. By satisfying these conditions, economic experiments ensure incentive compatibility and thus participants' true preferences towards

---

<sup>1</sup>Some exceptions exist (e.g., Carrion & Levinson, 2013; Fezzi et al., 2014; Prato et al., 2014), though historically RP data is rarely used to estimate VTT and VOR because many contexts cannot be (easily) examined in a real-world driving study (e.g., certain roads may not yet exist, certain toll levels may not yet exist). More recently, there has been a renewed interest in using RP data for VTT research (e.g., Varela et al., 2018).

an attribute via the experiment (for designing economic experiments in transportation see Dixit et al., 2017).

In this paper, we test the hypothesis that—in line with the literature— an incentive-compatible experiment may reveal different route choice behaviour than a typical SCE. We use two data collection techniques in a hypothetical binary route choice network, namely (i) a typical SCE in an online survey, and (ii) an economic driving simulator experiment (DSE). In contrast to the typical SCE, the economic DSE requires drivers to experience the consequence (through simulation) in three ways: the travel time of the chosen route; the unreliability of that travel time through the presence of an alternative with an unknown travel time; and a financial consequence due to the deduction of toll costs incurred from the initial endowment allocated to respondents. In this study, a within-subject treatment is used wherein all respondents complete both (i) the SCE, and (ii) the economic DSE, where the choice tasks in both elements of the study are identical.

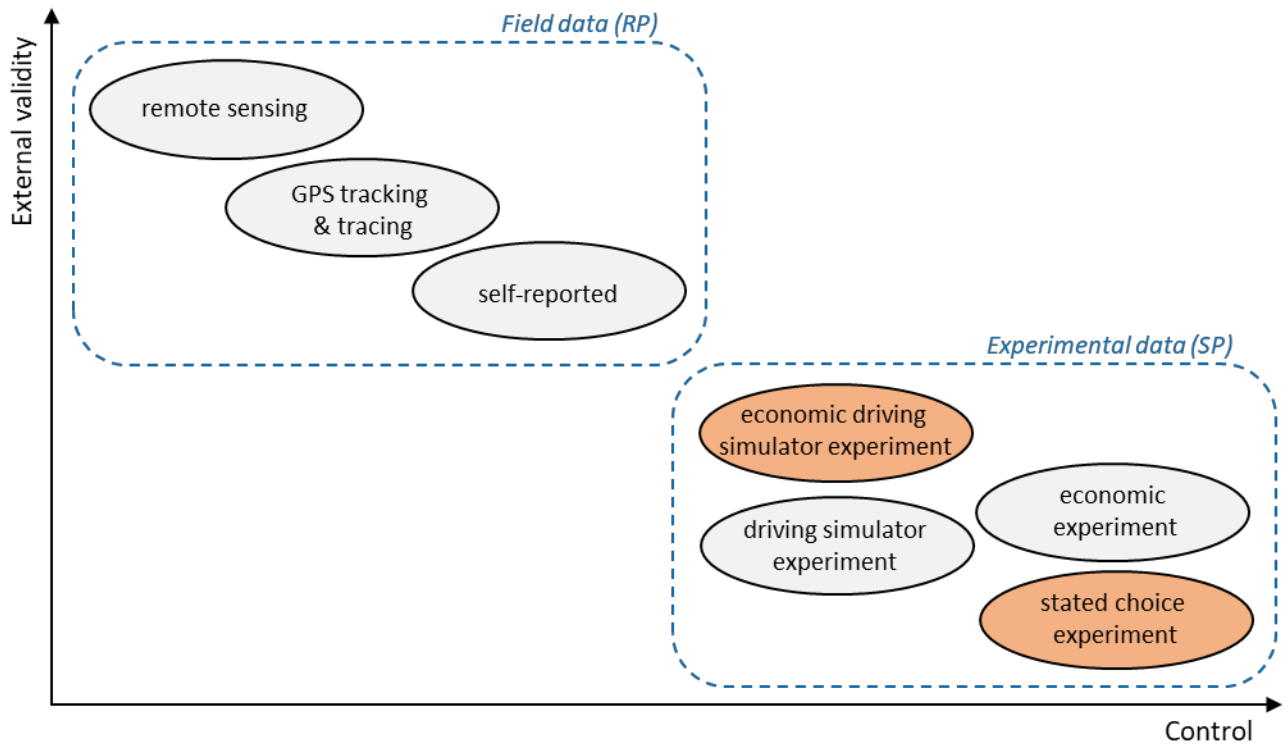
Whereas SCEs have been widely used to study route choice behaviour, DSEs have been mainly used to investigate operational driving behaviour (car following, lane changing) and safety, and less frequently to study route choice behaviour. To the best of our knowledge, this is the first study of its kind and this paper makes an additional contribution to the existing literature by being the first to use a DSE to estimate VTT and VOR measures in a route choice context.

The remainder of the paper is organised as follows. First, we provide a literature review of data collection techniques in a route choice context and their features, followed by a description of the methodology and data collection for the present study. Then, we present our data analysis, which includes our econometric framework for model estimation, and discussion of results. Finally, we draw conclusions with respect to the impact of economic incentives and simulated experiences in the context of route choice experiments.

## 2. Route choice data collection techniques and their features

Given the vast amount of literature on the collection of choice data, we limit this literature review to data collection techniques for car drivers in a route choice context. We discuss these data collection methods in the light of the trade-off between (external) validity (do the data reflect real behaviour) versus the degree of experimental control in collecting the data, illustrated in Figure 1, in which field data (RP) and experimental data (SP) are placed along both these axes. With respect to field data, the analyst has little control over the environment in which the data is collected, but the data has high external validity (i.e., is realistic thus exhibiting low hypothetical bias). In contrast, while experimental data has lower external validity, it can be collected with a high level of experimental control.

We will now look at each approach in more detail. First, we focus on RP methods and discuss their limitations, with respect to our research question, in Section 2.1. Then, we focus on SP data collection techniques in Section 2.2 and present the rationale for using an economic driving simulator experiment in our study.



**Figure 1** Route choice data collection techniques

## 2.1 Revealed preference

RP data includes drivers' route choices in a real-world setting. Such data can be divided into two categories, 1) self-reported questionnaires and interviews (drivers complete a questionnaire regarding their past route choices), and 2) GPS experiments or remote sensing (drivers are observed in real world traffic). These two categories result in the collection of travel time data that is either a subjective self-reported measurement or an objective observed measurement (Carrion & Levinson, 2012).

Several studies have used objective travel time observations from smartphones and other GPS devices in order to estimate route choice models for car drivers. Some of these route choice models were estimated to examine route switching behaviour. For example, van Essen et al. (2019) examined route choice behaviour from data collected via a real-world experiment (research vehicles fitted with GPS devices) and found that travel time information provision reduces route switching tendency and also helps in selecting the shortest route. Vacca et al. (2019) examined factors influencing route switching behaviour of drivers in Italy using GPS data from mobile phones and discovered that travel habits (i.e., a higher number of times a route is selected) likely to decrease route switching propensity. Papinski et al. (2009) explored reasons for different route selections and route switching behaviour in Canada. There are a number of studies that use GPS data to examine route switching behaviour (see Djukic et al., 2016; Ramos et al., 2012). Some studies use remote sensing (a data collection technique that observes vehicles without drivers being aware of it) to study route choice behaviour. For instance, Knoop et al. (2009) find route trajectories using video. Other studies used mobile positioning data to examine drivers' route choices, for example Ahas et al. (2010) used mobile positing data to examine commuter route choice patterns over weekdays and weekends. Mobile phone location data is another emerging area of research, and lends itself especially to modelling route choice for longer distance travel, where Bwambale et al. (2019a) show how such data can also be used to estimate meaningful VTT measures.



There exists only a very limited number of studies that have estimated VOR using objective travel time measurement. Carrion & Levinson (2013) estimated VOR using travel time data collected from GPS devices. Prato et al. (2014) estimated VOR from a GPS devices data in Denmark and recommend exploiting GPS data as technology is getting cheaper and the use thereof could result in “real” large scale models. In work related to travel time variability, Bwambale et al. (2019b) show how mobile phone data can be used to model departure time choices.

All of the above studies use objective travel time measurement; however, such studies (particularly when mobile positing data or GPS tracking are used) may not necessarily provide any situational characteristics (e.g., route descriptions, participants’ movements, attitudes, frequently visited locations) that can be collected in traditional questionnaire (Rieser-Schüssler & Axhausen, 2014). On the other hand, subjective travel time measurements extracted via questionnaires and interviews may reflect drivers’ post-journey perception of travel time of a route rather than the actual travel time. To estimate VOR from subjective travel time measurement, there is one study (to the best of our knowledge) by Brownstone & Small (2005) that uses self-reported RP data.

There are also a few studies which investigate drivers’ subjective perception of travel time to what they actually experience. Peer (2013) evaluated perception errors of drivers between objective travel time measurement (travel time obtained from cameras) and subjective travel time measurement (travel time reported by drivers) for morning commuters and found that there is a perception error between these two travel times as drivers on average overstate the reported travel time by a factor of approximately 1.5 compared to observed travel time. Carrion (2013) investigated causes and consequences of travel time perception errors for drivers’ work trips (where travel time data is collected using GPS devices and self-reported questionnaires) and found that drivers’ route choices are impacted by perceptions errors and cognitive biases, and recommended that transport planning models should instead consider drivers perception of travel time distributions. Of course, a counter-argument is that if the drivers actually use the perceived travel times when making their travel decisions, then this is what should be used in modelling too.

In summary, the main advantage of RP data (using any source of travel time measurement) is that it contains the actual choice of drivers observed in the real world (assuming truthful reporting in the case of questionnaires and interviews) and hence hypothetical and strategic bias is practically eliminated. However, there also exist disadvantages of RP data. First, one can only examine existing contexts, e.g. if tolls do not exist one cannot examine the impact of tolls. Secondly, attribute levels may not exhibit sufficient variation, e.g. it is unlikely that toll levels can be varied on an actual road for the purposes of experimentation. Thirdly, while the chosen alternative is observed, the other alternatives in the consideration set are not easily observed<sup>2</sup>. Fourthly, there is very little experimental control, e.g. the analyst cannot control traffic conditions. For these reasons, most route choice studies are conducted in a hypothetical context.

## 2.2 *Stated preference*

In SP studies, participants are asked to make choices in a series of hypothetical choice tasks that allow the use of alternatives, attributes, and attribute levels that cannot be observed in the current market and provides the analyst full control over the information presented to respondents. SP data collection techniques are typically referred to as experiments, and several classes can be distinguished, namely stated choice experiments, driving simulator experiments, and economic experiments. In this paper, we also consider the combination of the latter two. These types of experiments are discussed in the following subsections.

---

<sup>2</sup> Unchosen alternatives can be constructed using a range of methods but can be complex, take time and thus come at a cost, and are also subject to error.

### 2.2.1 Stated choice experiments

Stated choice experiments (SCEs) were originally proposed in the consumer research field (Louviere & Hensher, 1983) and since the 1990s, have received considerable attention in the field of route choice behaviour. These experiments are often used to collect data for the estimation of the monetary value of travel time changes, i.e. VTT, and changes in travel time unreliability, i.e. VOR (Hensher, 2001; Senna, 1994; Small, 1999), and investigating the factors influencing route choices (Abdel-Aty et al., 1997). Participants of these experiments typically provide responses via an online survey, a pen-and-paper survey, or to a personal interviewer that brings a laptop to the participants' home or workplace. SCEs can also be conducted in a computer laboratory which allows analysts to show more complex choice tasks and could potentially also collect eye-tracking data, or other extra behavioural/processing data. Examples of studies that use SCE to examine route choices include Ben-Elia et al. (2013a) and Razo & Gao (2013), who examine the effect of information provision on driver's route choice behaviour via advanced traffic information systems (ATIS). More recent examples related to ATIS include Qi et al. (2019) and Meneguzzer (2019).

Notwithstanding the very high experimental control of SCE, it is often argued in the literature that the SCE responses provided by participants tend to deviate from their responses when faced with the same situations in a real-life setting, with the discrepancy often referred to as "hypothetical bias". Several studies have compared outcomes from SCE data with RP data. Fifer et al. (2014) analysed RP data (collected by recruiting drivers in 10-weeks real-world driving study using GPS devices) and data from a SCE (by the same drivers) and found a difference between drivers' choices. In addition, studies reported that the VTT and VOR are underestimated in SCEs compared to RP data (Brownstone & Small, 2005; Isacsson, 2007), while others reported higher values in SCEs than RP (Wardman & Whelan, 2001). According to Fifer et al. (2014), hypothetical bias is a significant concern in SCEs. However, there is often no alternative to SCEs if we are interested in determining preferences towards, e.g., a newly proposed toll road or a newly proposed public transport service for which no RP data can be collected. There are various studies which investigate the issue of hypothetical bias and how to reduce it (Beck et al., 2016; Hensher, 2010).

### 2.2.2 Driving simulator experiments

Another technique to collect route choice behaviour data, in a hypothetical context, is to use driving simulators wherein the attributes of a route (e.g., travel time) can be experienced virtually. Driving simulator experiments (DSEs) allow the analyst more control than field experiments (including simulating alternatives, attributes and attribute levels that cannot yet be observed on existing roads) and are generally cheaper to conduct than real-world driving studies. Driving simulators have been used extensively in road safety research (e.g. Donmez et al., 2007; Young et al., 2014), and increasingly also to study other aspects of traffic operations and driving behaviour, such as hysteresis (e.g. Saifuzzaman et al., 2017); drivers' responses to advanced traveller information systems (e.g., Bonsall, 2004; Koutsopoulos et al., 1994); and interactions with connected and / or automated vehicles (e.g. Ali et al., 2020; Sharma et al., 2019), to name just a few examples. Bonsall (2004) argued that route choices in driving simulators provide more reliable data compared to data collected via typical surveys, the main advantage is that there is no need to inform a driver about an attribute of a route since the driver can experience it in the simulator.

In the past two decades, driving simulators have been increasingly used to investigate a range of driving and travel behaviour, including route choice behaviour. Bonsall et al. (1997) compared drivers' route choices in a driving simulator to their real-life decisions and found the decisions identical. Hess et al. (2020b) compare drivers' lane changing behaviour in an SCE and DSE and conclude that there are similarities as well as differences and suggests combining both data types. Despite the fact that DSEs allow participants to experience route attributes in a reasonably realistic virtual environment, they may still suffer from hypothetical bias. In real-life, driving is a goal-oriented activity where driving

is for a reason (Levinson et al., 2004), which is why route choices made by a participant in a simulation may not reveal the participant's true preferences towards an attribute (e.g. travel time). For instance, arriving too late at a destination in a simulation is not the same as arriving too late for a meeting in real life.

An alternative to driving simulators is the use of virtual reality. For instance, Blissing et al. (2019) and Branzi et al. (2017) have conducted virtual reality experiments to study driving behaviour but did not examine route choice behaviour.

### 2.2.3 Economic experiments

In economic experiments, the participants' reward structure (incentive) is designed using principles of induced value theory (Smith, 1976) which ensures incentive compatibility meaning that conditions for internal and external validity are satisfied. This differentiates typical SCEs from economic experiments. In a route choice context, most of the economic experiments<sup>3</sup> are conducted to study the effect of information provision (Ben-Elia et al., 2008; Ben-Elia et al., 2013b; Ben-Elia & Shiftan, 2010) in which participants' monetary incentive payments are dependent on the choices made during the experiment. In their study, drivers were divided into two groups, a control group where no information regarding travel time distribution was provided, and a treatment group that received such information. Drivers were asked to select a make a selection in a binary route choice task, where one route was faster on average while the second route has a risky travel time (and could have a low or high travel time). They found that information provision can be useful in case of non-recurrent congestion, particularly when drivers have limited knowledge of the road network.

In economic experiments, participants typically complete the choice tasks sitting in front of a computer in a laboratory environment. These types of experiments can involve either an experiment where the payoffs are dependent only on the choices made by the respondent alone, or experiments where the payoffs are also dependent on the choices made by other participants<sup>4</sup> (commonly referred to as an economic game). In cases where such interactive decision making is present, in order to reach an equilibrium state, participants are typically faced with a relatively large number of choice tasks (e.g. in one of the studies reference above, one treatment was repeated 100 times for each participant).

### 2.2.4 Economic driving simulator experiments

In this paper, we rely on a type of economic DSEs conducted within the context of a driving simulator experiment. The advantage of conducting such an experiment over a typical economic experiment conducted in front of a computer in a laboratory, at least in the context of route choice, is that not only are the consequences of simulated choices felt via the passage of time, but respondents must also engage in the task of driving the chosen route and in this study, face the consequence of risk (travel time variability) and cost (paying for use of a toll road). This creates an environment that is closer to real-world driving and thus arguably enhances the attentiveness of participants to the nature of the choices and the outcomes thereof.

In a route choice context, there is only one published study (Dixit et al., 2015) which carried out an economic experiment with driving simulators (and lottery choice tasks as is common in economic

---

<sup>3</sup> Although in the literature, laboratory experiments and economic experiments are often used interchangeably, in this paper we differentiate between these two. We define laboratory experiments where conditions for induced value theory are not satisfied while in case of economic experiments, these conditions are fulfilled.

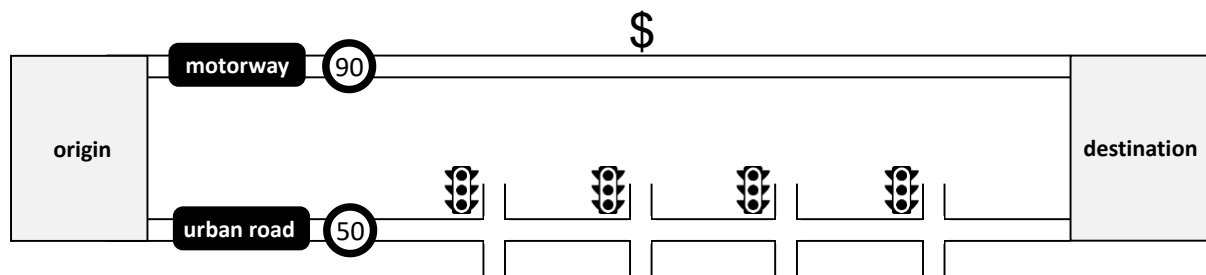
<sup>4</sup> In these interactive experiments, a participant maximises his/her utility by accounting for the influence of decisions by other participants through a payoff function. Such experiments are designed using concepts from game theory and the aim is to find an equilibrium solution (e.g., congestion games (Rosenthal, 1973)). The literature review in this paper does not discuss interactive experiments.

experiments), where the aim was to compare risk attitudes of two groups of drivers, namely general population vs. students<sup>5</sup>.

### 3. Study design

Our study consists of both an economic DSE and a typical SCE using a within-subject design, where the aim is to compare route choice behaviour of general population respondents in the two data collection techniques, and where the choice tasks are identical with the exception of simulated experiences and monetary incentives.

In this study, we focus on a hypothetical binary route choice network as shown in Figure 2. Participants are asked to imagine travelling from home to work by car with two possible route alternatives: Motorway and Urban Road. Each participant faces multiple choice tasks. The motorway alternative with a speed limit of 90 km/h and no traffic lights is designed to have a reliable travel time of 6 minutes across all choice tasks (i.e. exhibits no travel time variation), but the toll cost varies between \$1 and \$3<sup>6</sup>. The urban road alternative has a speed limit of 50 km/h and whilst having no toll, a driver will encounter four traffic lights that make the travel time unreliable, where it varies between 4 and 12 minutes according to a given probability distribution that changes across choice tasks.



**Figure 2 Hypothetical network**

In Section 3.1, we discuss the experimental procedure and in Section 3.2, we describe the basic route choice model that we aim to estimate. Then in Sections 3.3 and 3.4, we describe the stated choice and driving simulator experiments, respectively, used to collect the data to estimate this route choice model. The experimental design that is underlying both experiments is presented in Section 3.5.

#### 3.1 Experimental procedure

Participants completed the study in two phases: (i) a typical SCE consisting of five hypothetical choice tasks via an online survey, and (ii) the *same* five choice tasks where participants were made to experience their choices in the driving simulator (DSE) in the laboratory. To control for order effects, we divided participants into two groups. In one group, participants completed the SCE before the DSE (order 1), while participants in the other group first completed the DSE (order 2). Participants were not told that both experiments contain the same choice tasks.

The main objective of our study is to compare outcomes between SCE and DSE and using the same choice tasks and participants (i.e., a within-respondent setup) rules out that any differences in

<sup>5</sup> The purpose of this study was to determine if the risk attitudes of students are similar (or different) to the general population and whether students could be used instead of the general population in these experiments (since students are relatively inexpensive to recruit). Though they utilised a within-subject design as well, the two options (safe and risky) in the lottery choice tasks are framed in terms of lottery prizes (and probabilities) while in the driving simulator task, this is done via traffic congestion.

<sup>6</sup> Australian dollars, where 1 Australian dollar is 0.79 US dollar or 0.64 euro (price level 28 Feb 2018, the approximate time period around which the experiment was conducted).

outcomes could be attributed to differences in choice tasks or differences in participants, ensuring maximising statistical power in hypothesis testing. While we believe it is unlikely for respondents to remember the exact choice tasks, we acknowledge that this setup has the potential risk that participants do remember both the choice tasks and their choices. In order to minimise such risk, a minimum period of two weeks was given between the two experiments for each participant.

### 3.2 Stated choice experiment

Presenting travel time unreliability to participants in a SCE is not a straightforward exercise as the presentation formats vary considerably across empirical studies. For instance, Black & Towriss (1993) (cited in Tseng et al., 2009) suggested that participants can interpret a 5-point travel time distribution (in contrast to a 10-point) well. Tseng et al. (2009) conducted face-to-face interviews with 30 participants wherein they presented eight different formats taken from empirical studies in the literature on travel time unreliability. Based on responses to several indicators (e.g., clarity of reliability presentation, how easy it is to make a choice between two alternatives etc.) they recommended using a verbal description instead of a graph showing a probability distribution.

For this study, we designed the format of SCE choice tasks by considering the recommendations of Tseng et al. (2009). An example of the SCE choice task presented to participants is shown in Figure 3. We explained to the participants that only the travel times of the urban route and the toll cost of the motorway vary over choice tasks, indicated in red.

Motorway	Urban road
Speed limit of <b>90 km/h</b> , <b>no traffic lights</b> .	Speed limit of <b>50 km/h</b> , <b>four traffic lights</b> .
The travel time is <b>6 minutes</b> every day.  <div> <div>6</div> <div>6</div> <div>6</div> <div>6</div> <div>6</div> </div>	The travel time varies. You will experience one of the following <b>travel times (in minutes)</b> with equal probability:  <div> <div>8</div> <div>10</div> <div>10</div> <div>10</div> <div>12</div> </div>
Toll cost: <b>\$ 2.00</b>	Toll cost: <b>\$ 0.00</b>

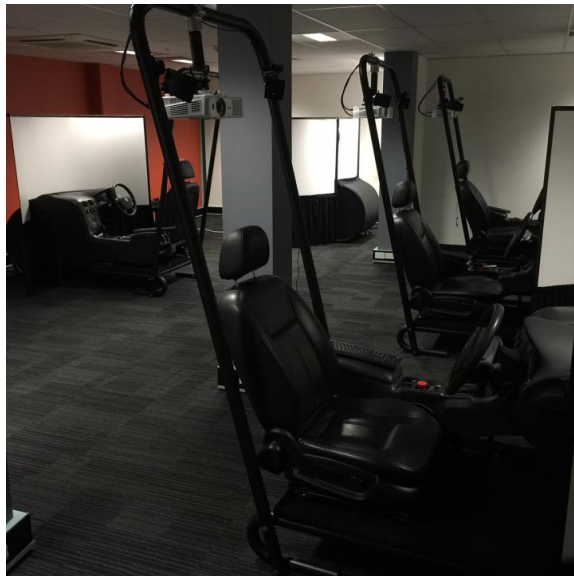
**Figure 3** Example of one of the SCE choice tasks presented to participants

### 3.3 Driving simulator experiment

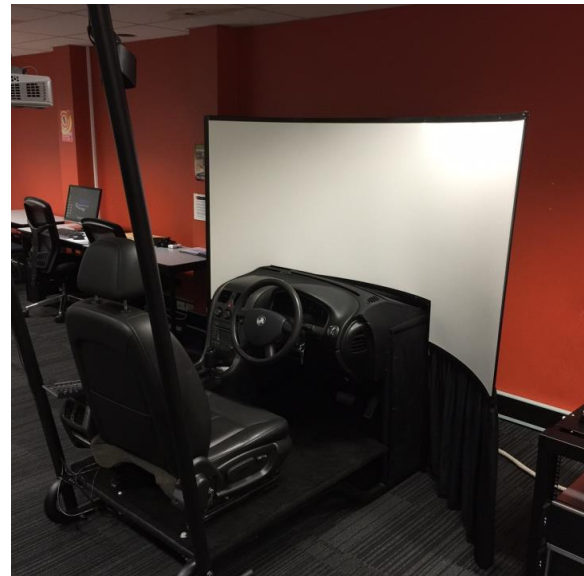
The driving simulators used in this study are based at the Travel Choice Simulation Laboratory at The University of Sydney (TRACSLab@USyd), as shown in Figure 4. The lab consists of five driving simulators which allow the analyst control over the driving environment (this includes the road system, traffic signals and controls, as well as computer-simulated cars as background traffic) and facilitate the capture of all decisions made by a maximum of five human drivers at the same time. The driving simulators are detached and transformed Holden Commodores. Each simulator is comprised of functional pedals, steering wheel, automatic gearbox (neutral, drive, reverse, etc.), indicator levers, dashboard, radio/CD player and automatic/motorized seat adjustment controls. To run a simulation on the driving simulators and to control traffic, the SimCreator software (developed by Realtime Technologies Inc.) is used. For the purposes of this experiment, driving in the simulator was restricted



to 12 minutes on any one route in order to avoid simulator sickness, with breaks of at least three minutes between consecutive driving episodes.



**(a) Overview of the TRACSLab@USyd**



**(b) Cockpit of the simulator**

***Figure 4 Driving simulators at TRACSLab@USyd***

In the driving simulator, the motorway route has no intersections or buildings, thereby providing an experience of driving on a highway, as shown in Figure 5(a). On the other hand, the urban road alternative replicates characteristics of a usual urban road which consists of a number of buildings, intersections, and traffic lights, see Figure 5(b). Both routes also have speed signs and other road signs. Travel times on the urban road are set by adjusting the traffic light settings. Furthermore, we control the speed of the vehicles by generating head and tail vehicles that drive exactly at 50 or 90 km/h for the urban road and motorway, respectively. Before choices are made in the simulator, each participant is asked to conduct a test drive on both the motorway and the urban road.

The DSE is the same as the SCE except for the fact that in the DSE, the participant is faced with simulated experiences of the route travel time, travel cost, and travel time unreliability after having made a route choice decision.



**(a) Motorway**



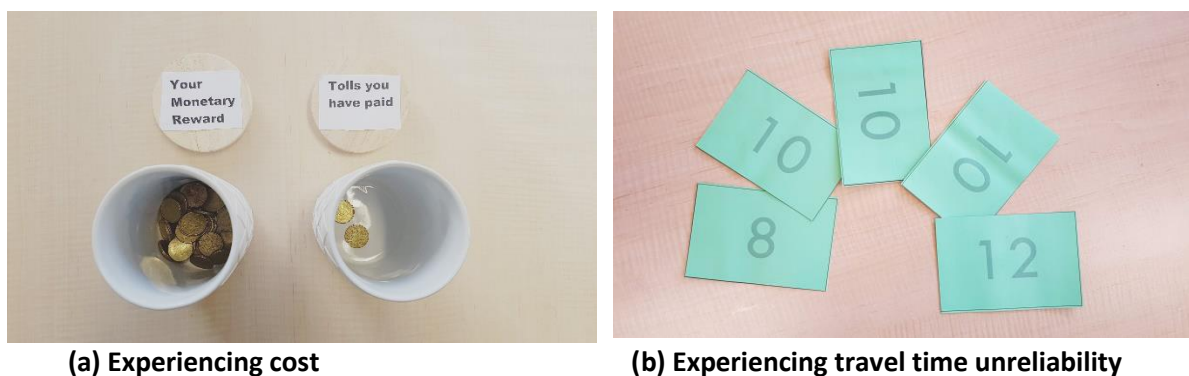
**(b) Urban Road**

***Figure 5 Routes in driving simulator***

### 3.3.1 Experiencing cost

Participants are told during recruitment that their participation reward will vary between \$40 and \$60 and that in order to be eligible, they must complete all phases of the study<sup>7</sup>. Adopting principles from experimental economics, their final reward was dependent on their route choices, which involves paying tolls costs when selecting the motorway alternative in the DSE.

Each participant receives an initial endowment of \$60, shown to the participant in a jar with one-dollar coins; see Figure 6(a). To illustrate further, suppose that the participant is faced with the choice task shown in Figure 3. If the participant selects the motorway alternative, \$2 is deducted from their initial endowment by transferring two coins in front of the participant from the jar in such a way that the respondent can clearly see the payment being made. After the experiment, the participant retains any remaining reward (paid out via a gift card). If the participant selects the urban road alternative in a given task, which has no toll, no money is deducted from their endowment in that task.



**Figure 6 Simulated experiences**

### 3.3.2 Experiencing travel time and unreliability

Suppose again that the participant is faced with the choice task shown in Figure 3. If the participant selects the motorway alternative, the participant is then asked to drive this route (with a travel time of 6 minutes) in the simulator and \$2 is deducted from their endowment. If the participant chooses the urban road alternative instead, he or she is asked to randomly select one card from a set of five cards (blind to the participant) where each card corresponds to one of the travel times in the distribution given in the choice task. This means that in case of the choice task in Figure 3, there is one card showing 8 minutes, three cards card showing 10 minutes, and one card showing 12 minutes (cf. Figure 6(b)). The analyst changes the settings of the driving simulator to reflect the travel time that was randomly selected by the participant, and the participant then drives on that route and experiences the travel time. It should be noted that respondents were not told the travel time that had been selected (as no driver knows with confidence what travel time they will experience prior to departure).

Following the driving task, all five cards with travel times, including the selected one, are revealed to the participant such that the participant can verify the travel time distribution matched the one presented in the choice task and that the selected travel time was the one that they experienced. It is important that the analyst does not reveal the travel time selected by the participant until after the drive in the simulator in order to ensure that the participant experiences travel time unreliability. The

<sup>7</sup> Participants received a fee of A\$15 if they were not able to complete the study, e.g. due to motion sickness, or if they did not adhere to the rules regarding speeding and driving off-road in the driving simulator.

verification that the card selected by the respondent was the travel time experienced in the simulator, and that the distribution from which the card was selected was the same as the distribution presented in the choice task, is important in order to maintain trust between the respondent and the analyst, and in the integrity of the experimental procedure.

### 3.4 Experimental design

We define four levels for toll cost, namely \$1, \$2, or \$3 for the motorway alternative and \$0 for the urban road alternative. Regarding the travel time distributions, the motorway alternative has a fixed travel time of 6 minutes (which equates to a mean travel time of 6 and a variance of 0) while for the urban road alternative, we consider combinations of five travel times, where each travel time can be 4, 6, 8, 10, or 12 minutes. These attribute levels were chosen for several reasons, but an important consideration is that each participant was limited to 90 minutes in the simulator. If a respondent was unlucky enough to generate 5 choice tasks each with a travel time of 12 minutes, this would require 60 minutes of driving time, with a further 15 minutes of enforced breaks (to avoid simulator sickness). With other associated time requirements of the experiment, this upper value of 12 minutes per task was the maximum that could reasonably be expected to be completed within the timeframe given. With regards to the remaining mix of times and costs, with these attribute levels, if a respondent was to select only the motorway option in each task, the expected cost across the 5 choice tasks would be \$10, with an approximate time saving of 30 minutes. This would equate to a value of time of approximately \$20 per hour, which is close to the VTT used by New South Wales Government in Australia.

In determining the correct design approach, it is imperative to already anticipate the specification of the econometric model, a point we turn to now. The two route alternatives are described by three attributes, namely travel (toll) cost, mean travel time, and the standard deviation of travel time (describing travel time unreliability). Adopting random utility theory in order to describe route choice, we define the utility of route  $i$  for respondent  $n$  in choice task  $t$ , denoted by  $U_{nti}$ , as

$$U_{nti} = V_{nti} + \varepsilon_{nti}, \quad (1)$$

where  $V_{nti}$  is the systematic utility and  $\varepsilon_{nti}$  is a randomly distributed error term. We consider two routes, and assume that each route is described by a given travel time distribution, denoted by  $\mathbf{T}_{nti}$ , and given toll costs,  $C_{nti}$ . The systematic route utility is assumed to depend linearly on the average travel time, standard deviation of travel time, and toll cost. Furthermore, we assume an alternative-specific constant for the motorway alternative, using a dummy coded variable  $M_i$  that equals 1 for the motorway route and 0 for the urban road. Therefore, systematic route utilities are defined as:

$$V_{nti} = \delta_M M_i + \beta_T E(\mathbf{T}_{nti}) + \beta_S \sqrt{\text{var}(\mathbf{T}_{nti})} + \beta_C C_{nti}, \quad (2)$$

where  $\delta_M$  is the alternative-specific constant for motorway and  $\boldsymbol{\beta} = (\beta_T, \beta_S, \beta_C)$  is a vector of marginal utility coefficients.

Eqn. (2) represents a mean-variance model that is often used to account for travel time unreliability. In the basic model, we assume that the error terms are independently and identically extreme value type I distributed, which means that route choice probabilities,  $P_{nti}$ , can be computed using a multinomial logit (MNL) model. That is:

$$P_{nti} = \frac{\exp(V_{nti})}{\sum_i \exp(V_{nti'})}. \quad (3)$$



There exist 3 possible profiles for the motorway alternative, given the fixed travel time and the three possible toll levels. On the other hand, with five travel times shown for the urban road alternative in each scenario, drawn from five possible values (4, 6, 8, 10 or 12 minutes), there exist  $5^5 = 3,125$  possible profiles for the urban road alternative. However, many of these profiles are essentially identical, as for example the travel time combination {4,6,6,10,12} represents the same distribution as {6,12,10,4,6}. Considering only unique travel time distributions, which we represent in increasing travel time order to participants, there are 126 profiles left for the urban road alternative. Additionally, we only consider travel time distributions with a unique mean and variance, e.g. {6,6,10,10,10} has the same mean and variance as {6,8,8,8,12}, which removes 44 profiles. This means that in total 246 different choice tasks (combinations of motorway and urban road profiles) exist.

Further, choice tasks where the urban road strictly or weakly dominates the motorway have been removed. For example, a choice task where the urban road has travel time distribution {4,4,4,4,4} is removed since the urban road strictly dominates the motorway (as it is cheaper, faster, and has the same reliability). If the urban road has travel time distribution {4,4,6,6,6} then it does not strictly dominate the motorway according to utility function (2) because while the urban road is cheaper and faster on average, it is less reliable. However, we can argue that a distribution of {4,4,6,6,6} is likely preferred over {6,6,6,6,6} despite travel time being less reliable, therefore we removed such choices tasks with a weakly dominant alternative from the candidate set. The final candidate set consists of 67 choice tasks.

Not all 67 choice tasks provide the same level of information for estimating the parameters in Eqn. (1). Given the limited number of participants in driving simulator experiments, we selected a subset of 15 choice tasks that provide high (Fisher) information for estimating the parameters in our model in Eqn. (1). This was achieved by generating a D-efficient experimental design through the modified Federov algorithm in Ngene (ChoiceMetrics, 2018) using our candidate set with 67 choice tasks. This design minimises the standard errors of the estimated model parameters and thereby minimises the required sample size (Bliemer et al., 2008; Rose & Bliemer, 2009). For the travel time and travel cost attributes, we assumed parameter priors based on estimates reported by Bliemer et al. (2017), who conducted a stated choice survey regarding route choice in Australia. For the standard deviation of travel time attribute, we assumed a prior consistent with a reliability ratio<sup>8</sup> of 0.85 (which is the average of the range 0.2 to 1.5 as reported in De Jong & Bliemer, 2015). The mean and standard deviation of travel time were computed for each of the 67 travel time distributions in the candidate set and once the experimental design was generated, they were converted back to travel time distributions consisting of five travel times.

Given that each participant could only spend a maximum of 90 minutes in the lab as stated in our ethics approval, each respondent was asked to complete only five choice tasks. Therefore, we blocked the design into three blocks of five choice tasks each where we aimed to have some degree of attribute level balance within each block (e.g. in each block the participant will face each of the toll levels at least once). The final experimental design is presented in Table1.

---

<sup>8</sup> Reliability ratio (RR) is defined as VOR divided by VTT (or marginal rate of substitution between travel time and standard deviation of travel time). Assuming a RR of 0.85, the resulting prior for standard deviation of travel time is equal to -0.122.

**Table 1 Choice tasks and blocks**

Block	Motorway		Urban Road	
	Travel times (min)	Toll (\$)	Travel times (min)	Toll (\$)
I	6, 6, 6, 6, 6*	2	4, 4, 10, 12, 12**	0
	6, 6, 6, 6, 6	1	8, 10, 10, 10, 12	0
	6, 6, 6, 6, 6	1	4, 6, 12, 12, 12	0
	6, 6, 6, 6, 6	2	4, 6, 6, 8, 10	0
	6, 6, 6, 6, 6	3	8, 10, 12, 12, 12	0
II	6, 6, 6, 6, 6	1	4, 6, 8, 10, 12	0
	6, 6, 6, 6, 6	1	8, 10, 12, 12, 12	0
	6, 6, 6, 6, 6	1	4, 4, 6, 8, 8	0
	6, 6, 6, 6, 6	2	4, 4, 4, 6, 12	0
	6, 6, 6, 6, 6	3	4, 6, 10, 12, 12	0
III	6, 6, 6, 6, 6	1	4, 6, 6, 8, 8	0
	6, 6, 6, 6, 6	2	8, 10, 10, 10, 12	0
	6, 6, 6, 6, 6	1	4, 6, 10, 12, 12	0
	6, 6, 6, 6, 6	2	8, 10, 12, 12, 12	0
	6, 6, 6, 6, 6	3	4, 6, 12, 12, 12	0

\* 6 minutes every day

\*\* 4 minutes two days per week, 10 minutes one day per week, and 12 minutes two days per week.

## 4. Data collection and preliminary analysis

### 4.1 Participants

While many DSEs are conducted with students only, we opted to sample from the general population in order to get more variation in our sample (in particular with respect to age and income). We are not necessarily concerned with obtaining a representative sample since we are not using the estimated VTT and VOR for appraisal purposes, but rather are mainly interested in differences in outcomes between the in SCE and DSE. Unlike SCEs, it is not easy to recruit non-student participants for the DSEs (despite a reward of up to A\$60) since they typically require a larger time commitment and require the respondent to physically attend the lab.

We used two approaches to recruit participants, using advertisements and contacting participants from previous experiments conducted at TRACSLab. In advertisements, we adopted a number of different tools including a University of Sydney (USYD) volunteer database for research studies, posting study details via USYD Business School official Facebook and Twitter pages and the TRACSLab website, flyers at the University (including coffee shops), and free local classified ads (via [www.gumtree.com.au](http://www.gumtree.com.au)).

During the recruitment process, participants complete a screening questionnaire which included supplementary questions designed to elicit information on their driving experience (i.e., how long they have been driving), age, sex, income, level of education, and occupation. The participants had to fulfil the following requirements:

- Hold a driving license;
- Drive at least 10 minutes, two or more days per week;
- Be at least 18 years old;
- Not suffer from motion sickness, vertigo, vestibular migraines or epilepsy;
- Not suffer from any other medical conditions that impact the ability to drive.

The recruitment process was repeated three times in total, with waves in September 2017 to December 2017; February 2018; and finally November 2018. Although in these three attempts, we were able to recruit more than 300 participants, only 76 participants actually turned up at the TRACSLab. With two participants experiencing motion sickness, we retained a final sample of 74 participants. Each participant completed 5 choice tasks in both SCE and DSE, resulting in a total of 740 choice observations.

## 4.2 Descriptive statistics

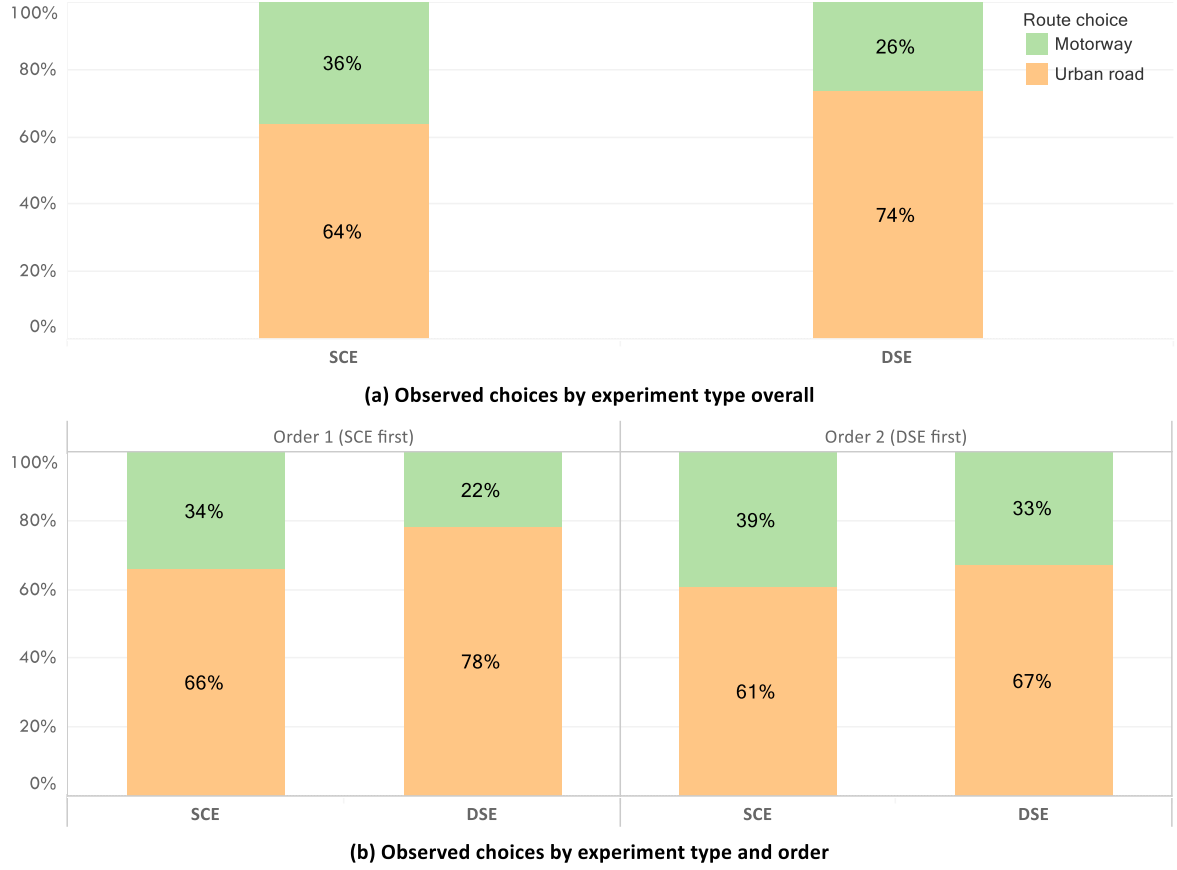
The socio-demographic characteristics of the 74 participants that completed the experiment are presented in Table 2. Comparing our sample to the population of Greater Sydney (ABS, 2016), we do see an over-representation of respondents who are more highly educated (university degree or higher), and aged between 30 to 39 years. Respondents in the sample also tend to have relatively higher incomes.

Regarding the order of experiments, 43 participants completed the two experiments in Order 1 (SCE first) and 31 participants in Order 2 (DSE first). Figure 7(a) shows the choice share of the motorway and urban road alternatives by experiment type. With respect to the experiment type, the Motorway is selected less often in the DSE (26%) compared to the SCE (36%). Therefore, these frequencies suggest that there may be differences in behaviour depending on the type of experiment.

Figure 7(b) further displays the choice shares depending on the order in which participants complete the experiments. We see that in Order 1 (SCE first), the motorway alternative is selected 34% in SCE compared to 22% in DSE. On the other hand, in Order 2 (DSE first), the motorway alternative is chosen 39% of the time in the SCE compared to 33% in the DSE. These differences in frequencies suggest that there exists an order effect, which is confirmed by a Chi-Square test ( $\chi^2(1, N = 74) = 5.66, p < .05$ ). We account for ordering effects in our econometric analysis in Section 5.

**Table 2 Socio-demographics of the participants**

Characteristics	Category	N=74	
		Sample (%)	Greater Sydney (%)
Gender	Male	44.6	49.3
	Female	55.4	50.7
Age (in years)	18-29	28.4	26.9
	30-39	35.1	23.9
	40-65	36.5	49.1
Annual Personal Income (A\$)	49,999 or less	29.7	50.4
	50,000-74,999	33.8	16.6
	75,000-99,999	17.6	10.5
	100,000 or more	16.1	13.4
	Not answered	2.7	9.2
Commuting to work (days per week)	$\geq 5$	44.6	NA
	4	21.6	NA
	$\leq 3$	33.8	NA
Occupation	Employed full-time	46.0	67.9
	Employed part-time	54.0	32.1
Education	Year 11 or less	0.0	17.2
	High School	13.5	35.5
	Associate degree (or Trade diploma)	14.9	15.2
	University degree or higher	70.2	22.5
	Not answered	1.3	9.6



**Figure 7 Observed route choices in SCE and DSE**

## 5. Model estimation and analysis

We adopted a structured and incremental approach toward estimating models. We first estimated our basic route choice multinomial logit (MNL) model which we outlined in the experimental design section (see Section 3.4). Next, we estimated heteroscedastic logit (HL) models to account for scale differences across the two experiment types (SCE and DSE experiments) and also due to order effects. Finally, we estimated a latent class (LC) model in order to account for preference heterogeneity across all attributes and the panel nature of the data (given that we have multiple observations from a single respondent). All models were implemented and estimated in R using Apollo (Hess & Palma, 2019). All models are estimated using pooled data (by combining SCE and DSE data). We excluded two participants from the dataset as they did not provide their income data, leading to a final sample of 72 respondents.

### 5.1 Multinomial logit models

In MNL 1, as discussed in Section 3.4, we estimate  $\delta_M$  and  $\beta$ . In MNL 2, we investigate whether preferences for the motorway alternative are different across the two types of experiment, by estimating an additional shift parameter  $\delta_{MD}$  for the motorway alternative in the DSE (i.e. when  $M_i = 1$ ) which is used when task  $t$  for person  $n$  is a simulator task (i.e. when  $D_{nt} = 1$ ):

$$V_{nti} = \delta_M M_i + \delta_{MD} M_i D_{nt} + \beta_T E(\mathbf{T}_{nti}) + \beta_S \sqrt{\text{var}(\mathbf{T}_{nti})} + \beta_C C_{nti}. \quad (4)$$

We next tested the inclusion of several sociodemographic variables where we found that only income had a statistically significant influence on route choice. In order to directly estimate the income elasticity  $\lambda_I$  we interact toll cost with a scaled income factor  $I_n$  for respondent  $n$ , where  $I_n$  is defined as income divided by average income. This gives us the following utility function for MNL 3:

$$V_{nti} = \delta_M M_i + \delta_{MD} M_i D_{nt} + \beta_T E(\mathbf{T}_{nti}) + \beta_S \sqrt{\text{var}(\mathbf{T}_{nti})} + \beta_C C_{nti} I_n^{\lambda_I}. \quad (5)$$

Table 3 presents parameter estimates for the three MNL models. In all models, we observe that the estimated parameters for the average travel time, toll cost, and travel time unreliability attributes have a negative sign, which is expected, and are statistically significant ( $p < 0.05$ ). In all cases, the parameter of the motorway dummy is negative, indicating that the (tolled) motorway option is less attractive than the (untolled but unreliable) urban road option (*ceteris paribus*).

In MNL 2 and MNL 3, we see that the additional DSE shift for the motorway constant is negative and statistically significant, implying that motorway is disliked more in DSE than SCE. A possible explanation is that participants are more averse to the tolled motorway (irrespective of the toll level) when they have to pay actual toll costs in the DSE. We investigate this further in the latent class model in Section 5.3.

In MNL 3, the estimate of  $\lambda_I$  shows a negative and statistically significant income elasticity towards toll cost, meaning that the sensitivity towards toll cost decreases with an increase in income. The relatively low value for the  $\lambda_I$  may be explained by the fact that we have mostly high income participants in our sample. The improvements from MNL 1 to MNL 2 (the loglikelihood ratio (LR) test statistic is equal to 9.72 while  $\chi^2_{1,95\%} = 3.84$ ) and then MNL 2 to MNL 3 (the LR test statistic is equal to 11.01 while  $\chi^2_{1,95\%} = 3.84$ ) are statistically significant. MNL 3 will serve as the starting point for estimating heteroscedastic models in the next section where we take possible scale differences into account.

**Table 3 Estimation results for MNL models**

	MNL 1	MNL 2	MNL 3
<i>Route attributes<sup>a</sup></i>			
Motorway ( $\delta_M$ )	-1.631 (-4.73)	-1.381 (-3.77)	-1.376 (-3.75)
Motorway x DSE ( $\delta_{MD}$ )	---	-0.543 (-2.49)	-0.551 (-2.49)
Avg. travel time ( $\beta_T$ )	-0.512 (-7.30)	-0.519 (-7.29)	-0.528 (-7.42)
St. dev. of travel time ( $\beta_S$ )	-0.274 (-4.28)	-0.278 (-4.29)	-0.278 (-4.24)
Toll cost ( $\beta_C$ )	-0.799 (-7.46)	-0.812 (-7.44)	-0.780 (-6.66)
Income elasticity ( $\lambda_I$ )	---	---	-0.322 (-2.09)
<i>Model fit</i>			
LL (0)	-499.07	-499.07	-499.07
LL (final)	-397.33	-392.467	-386.963
Adj. Rho sq.	0.196	0.204	0.213
No. of parameters	4	5	6
BIC	820.97	817.83	813.40

<sup>a</sup> Robust *t*-ratio values against zero are in brackets.

## 5.2 Heteroscedastic logit models

Our data originates from two different sources (SCE and DSE), and there may be different error variances (i.e. difference in scale) for each experiment type that need to be accounted for. In heteroscedastic logit (HL) models, we relax the assumption that the error variance is constant within the data. Furthermore, we also allow for difference in error variances based on order effects, i.e. which treatment was used first. This thus results in four scale terms in total.

We introduce dummy variable  $O_n$  that equals 1 if respondent  $n$  faces the two experiments in Order 1 (SCE first), and zero in case of Order 2 (DSE first). Applying different scale parameters for each type of experiment depending on the order leads to the following systematic route utilities:

$$V_{nti} = \bar{\mu}_{nt} \left( \delta_M M_i + \delta_{MD} M_i D_{nt} + \beta_T E(\mathbf{T}_{nti}) + \beta_S \sqrt{\text{var}(\mathbf{T}_{nti})} + \beta_C C_{nti} I_n^{\lambda_i} \right), \quad (6)$$

where we define

$$\bar{\mu}_{nt} = \begin{cases} \mu_{D1}, & \text{if } D_{nt} = 1 \text{ and } O_n = 1 \text{ (scale parameter for DSE, when SCE is taken first),} \\ \mu_{S1}, & \text{if } D_{nt} = 0 \text{ and } O_n = 1 \text{ (scale parameter for SCE, when SCE is taken first),} \\ \mu_{D2}, & \text{if } D_{nt} = 1 \text{ and } O_n = 0 \text{ (scale parameter for DSE, when DSE is taken first),} \\ \mu_{S2}, & \text{if } D_{nt} = 0 \text{ and } O_n = 0 \text{ (scale parameter for SCE, when DSE is taken first).} \end{cases} \quad (7)$$

In this model, which we refer to as HL1, we normalise  $\mu_{D1} = 1$  and estimate three scale parameters,  $\boldsymbol{\mu} = (\mu_{S1}, \mu_{D2}, \mu_{S2})$ . Results are presented in Table 4.

In HL 1 we observe that all estimates for  $\boldsymbol{\beta}$  remain statistically significant (and do so for all subsequent HL models). Examining the scale terms we note first that  $\mu_{S1}$  is not statistically different from the base  $\mu_{D1} = 1$ . Second, scale parameter  $\mu_{D2}$  is significantly different from the base of  $\mu_{D1} = 1$ , while  $\mu_{D2}$  is not significantly different from scale parameter  $\mu_{S2}$  (confirmed by testing whether  $\mu_{D2} - \mu_{S2} = 0$ , where this test incorporated the covariance between the estimates to ensure that the t-ratio on the difference reflects the maximum likelihood estimate properties of the original parameters; cf. Daly et al., 2012) and scale parameter  $\mu_{S2}$  is significantly different from 1 at the 10% level. Our findings thus suggest that scale differences exist by order but not by experiment type (SCE vs DSE).

Given the above results, a second model (HL 2) was estimated that only includes scale differences due to experiment order, such that the systematic route utilities simplify to:

$$V_{nti} = \mu_2^{1-O_n} \left( \delta_M M_i + \delta_{MD} M_i D_{nt} + \beta_T E(\mathbf{T}_{nti}) + \beta_S \sqrt{\text{var}(\mathbf{T}_{nti})} + \beta_C C_{nti} I_n^{\beta_i} \right), \quad (8)$$

where we estimate scale parameter  $\mu_2$  corresponding to Order 2 (DSE first,  $O_n = 0$ ) while scale is equal to 1 for Order 1 ( $O_n = 1$ ). Parameter estimates shown in Table 4 suggest a lower scale parameter, and hence more error variance, for data collected using Order 2. One *possible* explanation for this result is that those respondents who complete the DSE first may exhibit more variety seeking behaviour and try out both routes in the simulator, e.g. participants are likely to sample the Motorway a number of times, or they are still learning about the nature of the experiment. Then, when completing the SCE two weeks later, they may make similarly stochastic choices in the context of the absence of consequence to the choices made.

On the other hand, another *possible* explanation is that, in Order 1 (SCE first), participants may be observed to make relatively more consistent choices because all information is presented in the SCE to participants and no travel time, travel time unreliability, or toll costs are experienced. When confronted with the DSE that now includes experience/experimental consequence, respondents are

more incentivised to reveal their preferences towards these attributes and/or have learnt the nature of the experiment from the SCE and thus are less prone to variety seeking in the DSE.

While identifying one unique explanation for the result is not possible, a clear order effect can be observed via the impact on scale. Overall, we prefer HL 2 over HL 1 because model HL 2 is more parsimonious than HL 1 (the LR test statistic is equal to 0.71 whereas  $\chi^2_{2,95\%} = 5.99$ ).

Building on model HL 2, we further explored preference heterogeneity towards route attributes in the two experiment types by estimating multipliers in the case of DSE, which leads to the following utility functions for model HL 3:

$$V_{nii} = \mu_2^{1-O_n} \left( \delta_M M_i + \delta_{MD} M_i D_{nt} + \kappa_T^{D_{nt}} \beta_T E(\mathbf{T}_{nii}) + \kappa_S^{D_{nt}} \beta_S \sqrt{\text{var}(\mathbf{T}_{nii})} + \kappa_C^{D_{nt}} \beta_C C_{nii} I_n^{\beta_I} \right), \quad (9)$$

where we estimate preference parameters  $\beta$ , scale parameter  $\mu_2$  and attribute-specific multipliers  $\kappa = (\kappa_T, \kappa_S, \kappa_C)$ , which only apply in case of the driving simulator experiment where  $D_{nt} = 1$ . Table 4 presents parameters for HL 3, where we clearly observe that none of the multipliers are statistically different from 1. Initially, we expected that participants would be more sensitive to time and cost attributes in the DSE because they had to experience them. However, from HL 3 we cannot draw this conclusion. HL 3 also does not improve model fit over HL 2 (the LR test statistic is equal to 1.02 whereas  $\chi^2_{3,95\%} = 7.81$ ).

**Table 4 Estimation results for HL models**

	HL 1	HL 2	HL 3
<i>Route attributes<sup>a</sup></i>			
Motorway ( $\delta_M$ )	-1.977 (-3.34)	-1.784 (-4.14)	-1.578 (-3.27)
Motorway x DSE ( $\delta_{MD}$ )	-0.707 (-2.18)	-0.703 (-2.49)	-1.372 (-1.77)
Avg. travel time ( $\beta_T$ )	-0.721 (-5.56)	-0.654 (-9.84)	-0.606 (-6.77)
St. dev. of travel time ( $\beta_S$ )	-0.370 (-3.99)	-0.348 (-4.54)	-0.357 (-3.17)
Toll cost ( $\beta_C$ )	-1.010 (-6.07)	-0.937 (-7.54)	-0.990 (-4.78)
Income elasticity ( $\lambda_I$ )	-0.338 (-2.23)	-0.340 (-2.28)	-0.332 (-2.15)
<i>Multipliers<sup>b</sup> for DSE</i>			
Avg. travel time ( $\kappa_T$ )	---	---	1.244 (0.83)
St. dev. of travel time ( $\kappa_S$ )	---	---	1.045 (0.08)
Toll cost ( $\kappa_C$ )	---	---	0.910 (-0.28)
<i>Scale<sup>b</sup></i>			
Order 1 DSE ( $\mu_{D1}$ )	1.000	---	---
Order 1 SCE ( $\mu_{S1}$ )	0.847 (-0.69)	---	---
Order 2 DSE ( $\mu_{D2}$ )	0.523 (-2.94)	---	---
Order 2 SCE ( $\mu_{S2}$ )	0.582 (-1.87)	---	---
Order 1 ( $\mu_1$ )	---	1.000	1.000
Order 2 ( $\mu_2$ )	---	0.590 (-2.68)	0.568 (-2.81)
<i>Model fit</i>			
LL (0)	-499.07	-499.07	-499.07
LL (final)	-381.79	-382.14	-381.629
Adj. Rho sq.	0.217	0.220	0.215
No. of parameters	9	7	10
BIC	822.79	810.34	829.05

<sup>a</sup> Robust t-ratio values against zero are in brackets; <sup>b</sup> Robust t-ratio values against one are in brackets

### 5.3 Latent class model

To further explore potential differences in preferences, we estimate a latent class (LC) model with two classes,  $q \in \{1, 2\}$ , with the following class-specific systematic utility functions:

$$V_{niiq} = \mu_2^{1-O_n} \left( \delta_M M_i + \delta_{MDq} M_i D_{ni} + \kappa_T^{D_{ni}} \beta_{Tq} E(\mathbf{T}_{nii}) + \kappa_S^{D_{ni}} \beta_{Sq} \sqrt{\text{var}(\mathbf{T}_{nii})} + \kappa_C^{D_{ni}} \beta_C C_{nii} I_n^{\lambda_i} \right). \quad (10)$$

For the class assignment model, we simply use utility function  $V_q = \alpha_q$ , where we normalise  $\alpha_2 = 0$  and estimate only a constant for Class 1. We tried including sociodemographic variables in the class assignment utility function, but none were found to be statistically significant. We estimate class-specific preference parameters  $\delta_{MDq}$  and  $\beta_q$  while we keep  $\delta_M$  and  $\lambda_i$  generic across both classes since making them class-specific did not significantly improve the model fit. Also, attribute-specific multipliers  $\kappa$  are not considered class-specific in order to ensure that the model parameters are identifiable. We tried increasing the number of latent classes but a model with two latent classes was preferred based on the Bayesian Information Criterion (BIC), and consideration was also given to the application of the model and the practicability of the results (Beck et al., 2013).

Table 5 presents parameter estimates for the LC model where we accounted for the panel nature of the data (i.e., we observe multiple choice observations from the same participant). We see a substantial improvement in model fit compared to the MNL and HL models (the LR test statistic for HL 2 vs. LC model is equal to 73.10 whereas  $\chi^2_{8,95\%} = 15.51$ ). Based on  $\alpha_1 = 0.135$ , the shares of the two classes are 53% and 47% for Class 1 and Class 2, respectively, where these are not significantly different from an equal split.

Before discussing the two classes, we note that the order effect in terms of scale remains significant in this model. Further, we first discuss attribute-specific multipliers  $\kappa$ . All multipliers are less than 1 suggesting that the impact of these attributes is reduced in the DSE relative to the SCE, though  $\kappa_T$  and  $\kappa_S$  are both not significantly different from 1 indicating that statistically these attributes carry the same preference weight in both experiments. However, with the case of  $\kappa_C$ , the multiplier is significantly less than 1, indicating that the impact of the toll cost attribute itself is less in the DSE than in the SCE. While this result may seem counterintuitive, the outcome needs to be considered in parallel to the shift in the alternative-specific constant for the motorway ( $\delta_{MDq}$ ).

Looking at the class-specific parameters, we see two different patterns. The parameters for Class 1 seem to indicate that respondents belonging to this latent class have a large aversion towards the tolled motorway in the DSE as expressed by the large negative value for  $\delta_{MD1}$ , indicating that participants seem to prefer to avoid the motorway more in the DSE simply because it is a toll road, irrespective of the toll level (the estimated parameter for toll cost in the DSE is  $\kappa_C \beta_{C1} = -0.524$ ). We also observe a relatively high aversion to travel time variability.

In contrast, in Class 2, respondents do not seem to have such a strong aversion to the motorway in the DSE ( $\delta_{MD2}$  is not significant), nor to travel time variability ( $\beta_{S2}$  is not significant), rather seemingly trading mostly between travel time and cost. Toll levels matter to Class 2 respondents, given that the toll cost sensitivity is expressed by  $\kappa_C \beta_{C2} = -1.482$ , such that these respondents mainly switch to the urban road when the toll level on the motorway is high.

Overall, these results indicate that there are two classes of preference structures that exhibit differences in how the attributes of the alternatives are evaluated, and the significant cost multiplier  $\mu_C$  shows that the evaluation of the cost attribute, and thus ultimately the trade-off between cost and time, differs in the case of the DSE as compared to the SCE. We explore the implications of this result in the next section.



**Table 5 Estimation results for the LC model**

	Class 1	Class 2
<i>Route attributes <sup>a</sup></i>		
Motorway ( $\delta_M$ )	-0.639 (-1.02)	
Motorway x DSE ( $\delta_{MDq}$ )	-2.530 (-2.81)	-1.500 (-1.68)
Avg. travel time ( $\beta_{Tq}$ )	-0.926 (-6.34)	-1.013 (-3.77)
St. dev. Of travel time ( $\beta_{Sq}$ )	-0.678 (-3.10)	-0.306 (-1.45)
Toll cost ( $\beta_{Cq}$ )	-1.652 (-4.43)	-4.675 (-3.89)
Income elasticity ( $\lambda_I$ )	-0.237 (-1.66)	
<i>Multipliers <sup>b</sup> for DSE</i>		
Avg. travel time ( $\kappa_T$ )	0.757 (-1.57)	
St. dev. Of travel time ( $\kappa_S$ )	0.610 (-1.28)	
Toll cost ( $\kappa_C$ )	0.317 (-6.65)	
<i>Scale <sup>b</sup></i>		
Order 1 ( $\mu_1$ )	1.000	
Order 2 ( $\mu_2$ )	0.568 (-2.81)	
<i>Class allocation model <sup>a</sup></i>		
Class assignment factor ( $\alpha_q$ )	0.135 (0.38)	0.000
<i>Model fit</i>		
LL (0)	-499.07	
LL (final)	-345.591	
Adj. Rho sq.	0.276	
No. of parameters	15	
BIC	789.87	

<sup>a</sup> Robust t-ratio values against zero are in brackets; <sup>b</sup> Robust t-ratio values against one are in brackets

#### 5.4 Comparison of VTT and VOR

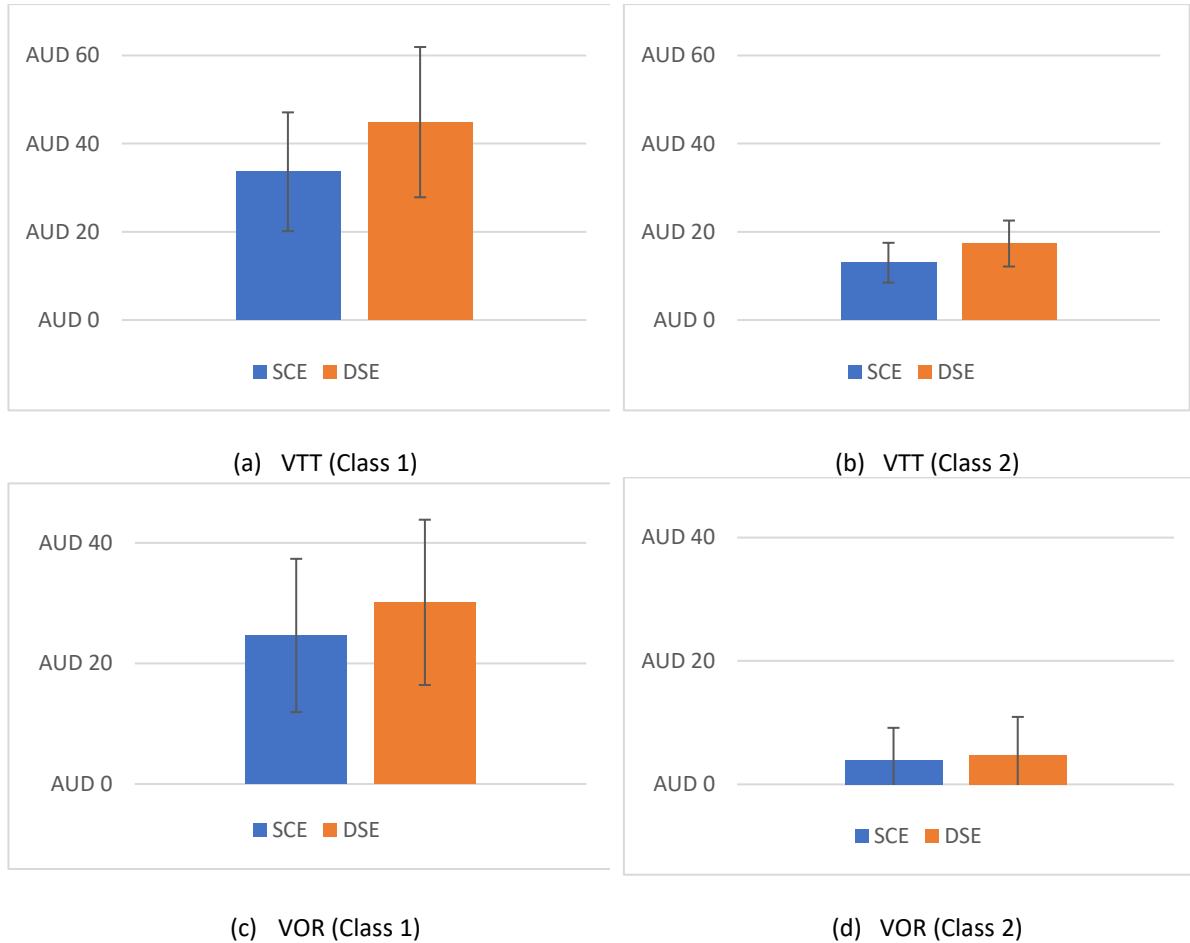
In Table 6, we summarise the VTT values ( $\beta_T / \beta_C$ ) and VOR values ( $\beta_S / \beta_C$ ), assuming average income, i.e.  $I_n = 1$ . We also report the reliability ratios (RR=VOR/VTT). In a recent study based in Sydney, Douglas & Jones (2018) discuss VOR using the reliability ratio for which they found an average value equal to 0.37. We observe that the VTT values in this experiment are relatively high compared to other VTT values estimated using population samples in Australia, in particular, A\$17.72/hr reported by TfNSW (2019), A\$17.39/hr estimated by Hensher (2019) and A\$7.33/hr estimated by Bliemer et al. (2017). Our high VTT value may be explained by two factors. First, most of our participants have a relatively high income (see Table 2, for descriptive statistics). Second, since the motorway in our study has a non-zero toll (whereas the Urban Road has zero tolls), there are some confounding effects between the motorway constant,  $\delta_M$ , and the toll cost attribute (with related parameter  $\beta_C$ ) in capturing the departure from a zero cost.

For the above reasons, the VTT and VOR values found in this study are not directly comparable to other values found in the literature. However, we remind the reader that the objective of this paper is not to generate VTT and VOR values that apply to the population, rather we are seeking to compare route choice behaviour in the two data collection techniques, and where the choice tasks are identical with the exception of simulated experiences and monetary incentives being enforced in the DSE.

**Table 6 Travel time-saving values (in A\$/hr)**

	MNL 1	MNL 2	MNL 3	HL 1	HL 2	HL 3		LC			
						SCE	DSE	SCE		DSE	
								Class 1	Class 2	Class 1	Class 2
VTT	38.41	38.35	40.59	42.85	41.88	36.69	43.10	33.63	13.00	44.88	17.35
VOR	20.57	20.54	21.34	21.99	22.29	21.62	23.14	24.64	3.93	30.13	4.81
RR	0.54	0.54	0.53	0.51	0.53	0.59	0.54	0.73	0.30	0.67	0.28

Figure 8 visualises VTT and VOR results from the LC model. The confidence intervals for the error bars are calculated via the Delta method which uses the robust variance-covariance matrix resulting from model estimation. It can also be seen that the VTT and VOR are statistically different across the classes, but not between SCE and DSE (within the same class). While we cannot conclude that VTT and VOR are different for SCE and DSE, we did observe significant differences in preferences towards toll road and toll level in the DSE. Therefore, failure to observe differences in VTT and VOR across experiment types may be the result of sample size limitations in this study.



**Figure 8 VTT and VOR (per hour) across the two classes in the LC model**

## 6. Discussion and conclusions

This paper investigates drivers' route choice behaviour in a typical SCE and a DSE that exhibits a greater degree of incentive compatibility. Initial analysis on the choice shares revealed that the motorway alternative was significantly less attractive in the DSE compared to the SCE, particularly when the SCE was completed first (Order 1). Subsequent modelling revealed that the Motorway alternative was indeed less attractive, as evidenced by the alternative specific constant which was significant across all estimated models; and the motorway alternative was found to be particularly less attractive in the DSE. Subsequent mathematical modelling revealed significant differences in the error variance of choices, with the consistently lower scale parameter for Order 2 indicating that choices exhibit greater error variance when the DSE is completed before the SCE.

Analysis also revealed preference heterogeneity among respondents; with one preference class broadly exhibiting strong aversion to the motorway alternative and travel time variability, while the other preference class showed dislike towards high toll cost. When examining any impact that the type of experiment might have on preferences, we found that the impact of cost was different in the DSE than in the SCE, in that the participants avoid the motorway more in the DSE. Regarding travel time and travel time unreliability, we could not find significant differences between the two types of experiments. Therefore, the type of experiment (with or without simulated experiences) seems to affect responses to monetary attributes differently to time attributes. Despite these differences, we could not reject the hypothesis that VTT and VOR are same in the typical SCE and the DSE, which is likely due to our limited sample size.

The similarities and differences between the typical SCE and incentive compatible DSE from this research may have implications for the design of route choice studies. Our study has some insights particularly when there is a tolled road, given that cost has a different impact in the DSE than in a typical SCE. Many decisions regarding new toll road projects are based on analysis from SP surveys, given the significant investments are made in constructing toll roads there is a need for robust inputs into this decision-making process. Assuming that the DSE is closer to the true preferences of a respondent through greater incentive compatibility, we recommend the adoption of incentive compatible DSE to allow for to the simulation of not only the new tolled route, but also to gather robust valuations of time and reliability that could be used to calibrate SCE data that may exhibit hypothetical bias. There is potential for future research to examine if the DSE experiments are better able to recover what would be revealed preferences. For route choice studies that include only travel time and travel time unreliability, a SCE may suffice.

Future research may also seek to provide a robust explanation for why the experiment order produces a significant difference in error variances estimated on the data, and in turn what this might mean for determining which process may produce more robust estimates of VTT and VOR. For example, if choices are relatively more deterministic if an SCE is completed first and is followed by a DSE (Order 1), then it may be the case that more robust values can be estimated cost effectively, by simply giving respondents a typical SCE and then replicating those choices a small number of times in a driving simulator as opposed to a complete replication of all choice tasks (for example, Hess et al. 2020b combined SCE and DSE data for efficiently estimating lane choice models).

Finally, there are several limitations to this study that should be noted. While the DSE simulates time and cost consequences for respondents, we could not include any outcome-related consequences such as being early/late for work. Also, it is possible that some people obtained a positive utility from driving in the simulator due to 'game' feeling (although all people were observed to be generally tired at the end of the DSE). Moreover, risk-seeking drivers may choose the urban road alternative because of the gambling aspect (although this may equally be true in real life route choices). As previously

stated, the sample size is small and not representative, therefore VTT and VOR values reported in this paper should only be used for model comparison and not for appraisal purposes. Lastly, for future research related to the experimental design of such experiments, we suggest adding a toll level of zero to the motorway to clearly disentangle the relationship between toll cost and travel time.

## References

- Abdel-Aty, M. A., Kitamura, R., & Jovanis, P. P. (1997). Using stated preference data for studying the effect of advanced traffic information on drivers' route choice. *Transportation Research Part C: Emerging Technologies*, 5(1), 39–50. [https://doi.org/10.1016/S0968-090X\(96\)00023-X](https://doi.org/10.1016/S0968-090X(96)00023-X)
- Abrantes, P. A. L., & Wardman, M. R. (2011). Meta-analysis of UK values of travel time: An update. *Transportation Research Part A*, 45, 1–17. <https://doi.org/10.1016/j.tra.2010.08.003>
- ABS. (2016). *Australian Bureau of Statistics, General Community Profile, Greater Sydney, 2016, data cube: Excel spreadsheet, viewed 10 June 2019,*.  
[https://quickstats.censusdata.abs.gov.au/census\\_services/getproduct/census/2016/communityprofile/1GSYD?opendocument](https://quickstats.censusdata.abs.gov.au/census_services/getproduct/census/2016/communityprofile/1GSYD?opendocument)
- Ahas, R., Aasa, A., Silm, S., & Tiru, M. (2010). Daily rhythms of suburban commuters' movements in the Tallinn metropolitan area: Case study with mobile positioning data. *Transportation Research Part C: Emerging Technologies*, 18(1), 45–54.  
<https://doi.org/10.1016/j.trc.2009.04.011>
- Ali, Y., Zheng, Z., Mazharul Haque, M., Yildirimoglu, M., & Washington, S. (2020). Understanding the discretionary lane-changing behaviour in the connected environment. *Accident Analysis & Prevention*, 137(January), 105463. <https://doi.org/10.1016/j.aap.2020.105463>
- Beck, M. J., Fifer, S., & Rose, J. M. (2016). Can you ever be certain? Reducing hypothetical bias in stated choice experiments via respondent reported choice certainty. *Transportation Research Part B: Methodological*, 89, 149–167. <https://doi.org/10.1016/j.trb.2016.04.004>
- Beck, M. J., Rose, J. M., & Hensher, D. A. (2013). Environmental attitudes and emissions charging: An example of policy implications for vehicle choice. *Transportation Research Part A: Policy and Practice*, 50, 171–182. <https://doi.org/10.1016/j.tra.2013.01.015>
- Ben-Elia, E., Di Pace, R., Bifulco, G. N., & Shiftan, Y. (2013a). The impact of travel information's accuracy on route-choice. *Transportation Research Part C: Emerging Technologies*, 26, 146–159. <https://doi.org/10.1016/j.trc.2012.07.001>
- Ben-Elia, E., Erev, I., & Shiftan, Y. (2008). The combined effect of information and experience on drivers' route-choice behavior. *Transportation*, 35(2), 165–177.  
<https://doi.org/10.1007/s11116-007-9143-7>
- Ben-Elia, E., Ishaq, R., & Shiftan, Y. (2013b). "If only I had taken the other road...": Regret, risk and reinforced learning in informed route-choice. *Transportation*, 40(2), 269–293.  
<https://doi.org/10.1007/s11116-012-9426-5>
- Ben-Elia, E., & Shiftan, Y. (2010). Which road do I take? A learning-based model of route-choice behavior with real-time information. *Transportation Research Part A: Policy and Practice*, 44(4), 249–264. <https://doi.org/10.1016/j.tra.2010.01.007>
- Black, I., & Towriss, J. (1993). *Demand effects of travel time reliability*. Center for Logistics and Transportation, Cranfield Institute of Technology.

- Bliemer, M. C. J., Rose, J. M., & Chorus, C. G. (2017). Detecting dominance in stated choice data and accounting for dominance-based scale differences in logit models. *Transportation Research Part B: Methodological*, 102, 83–104. <https://doi.org/10.1016/j.trb.2017.05.005>
- Bliemer, M. C. J., Rose, J. M., & Hess, S. (2008). Approximation of Bayesian efficiency in experimental choice designs. *Journal of Choice Modelling*, 1(1), 98–126. [https://doi.org/10.1016/S1755-5345\(13\)70024-1](https://doi.org/10.1016/S1755-5345(13)70024-1)
- Blissing, B., Bruzelius, F., & Eriksson, O. (2019). Driver behavior in mixed and virtual reality – A comparative study. *Transportation Research Part F: Traffic Psychology and Behaviour*, 61, 229–237. <https://doi.org/10.1016/j.trf.2017.08.005>
- Bonsall, P., Firmin, P., Anderson, M., Palmer, I., & Balmforth, P. (1997). Validating the results of a route choice simulator. *Transportation Research Part C: Emerging Technologies*, 5(6), 371–387. [https://doi.org/10.1016/S0968-090X\(98\)00003-5](https://doi.org/10.1016/S0968-090X(98)00003-5)
- Bonsall, P. W. (2004). Route Choice Simulators. In *Human Behaviour and Traffic Networks* (pp. 167–191). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-662-07809-9\\_7](https://doi.org/10.1007/978-3-662-07809-9_7)
- Branzi, V., Domenichini, L., & La Torre, F. (2017). Drivers' speed behaviour in real and simulated urban roads – A validation study. *Transportation Research Part F: Traffic Psychology and Behaviour*, 49, 1–17. <https://doi.org/10.1016/j.trf.2017.06.001>
- Brownstone, D., & Small, K. A. (2005). Valuing time and reliability: Assessing the evidence from road pricing demonstrations. *Transportation Research Part A: Policy and Practice*, 39(4 SPEC. ISS.), 279–293. <https://doi.org/10.1016/j.tra.2004.11.001>
- Bwambale, A., Choudhury, C. F., & Hess, S. (2019a). Modelling long-distance route choice using mobile phone call detail record data: a case study of Senegal. *Transportmetrica A: Transport Science*, 15(2), 1543–1568. <https://doi.org/10.1080/23249935.2019.1611970>
- Bwambale, A., Choudhury, C. F., & Hess, S. (2019b). Modelling departure time choice using mobile phone data. *Transportation Research Part A: Policy and Practice*, 130(October), 424–439. <https://doi.org/10.1016/j.tra.2019.09.054>
- Carrion, C. (2013). *Travel Time Perception Errors : Causes and Consequences* (Issue July). Retrieved from the University of Minnesota Digital Conservancy, <http://hdl.handle.net/11299/155609>.
- Carrion, C., & Levinson, D. (2012). Value of travel time reliability: A review of current evidence. *Transportation Research Part A: Policy and Practice*, 46(4), 720–741. <https://doi.org/10.1016/j.tra.2012.01.003>
- Carrion, C., & Levinson, D. (2013). Valuation of travel time reliability from a GPS-based experimental design. *Transportation Research Part C: Emerging Technologies*, 35, 305–323. <https://doi.org/10.1016/j.trc.2012.10.010>
- ChoiceMetrics. (2018). Ngene 1.2 User Manual & Reference Guide, Australia.
- Daly, A., Hess, S., & de Jong, G. (2012). Calculating errors for measures derived from choice modelling estimates. *Transportation Research Part B: Methodological*, 46(2), 333–341. <https://doi.org/10.1016/j.trb.2011.10.008>
- de Jong, G. C., & Bliemer, M. C. J. (2015). On including travel time reliability of road traffic in appraisal. *Transportation Research Part A: Policy and Practice*, 73, 80–95. <https://doi.org/10.1016/j.tra.2015.01.006>
- Ding, M., Grewal, R., & Liechty, J. (2005). Incentive-Aligned Conjoint Analysis. *Journal of Marketing Research*, 42(1), 67–82. <https://doi.org/10.1509/jmkr.42.1.67.56890>

- Dixit, V. V, Harb, R. C., Martínez-Correa, J., & Rutström, E. E. (2015). Measuring risk aversion to guide transportation policy: Contexts, incentives, and respondents. *Transportation Research Part A: Policy and Practice*, 80, 15–34. <https://doi.org/10.1016/j.tra.2015.07.002>
- Dixit, V. V, Ortmann, A., Rutström, E. E., & Ukkusuri, S. V. (2017). Experimental Economics and choice in transportation: Incentives and context. *Transportation Research Part C: Emerging Technologies*, 77, 161–184. <https://doi.org/10.1016/j.trc.2017.01.011>
- Djukic, T., Wilmlink, I., Jonkers, E., Snelder, M., & van Arem, B. (2016). *Exploratory Analysis of Traveller's Compliance with Smartphone Personal Route Advice: In a Field Trial Amsterdam*.
- Donmez, B., Boyle, L. N., & Lee, J. D. (2007). Safety implications of providing real-time feedback to distracted drivers. *Accident Analysis & Prevention*, 39(3), 581–590. <https://doi.org/10.1016/j.aap.2006.10.003>
- Douglas, N., & Jones, M. (2018). Estimating the Value of Private Travel Time for Nsw. *Australasian Transport Research Forum*, November, 1–31.
- Fezzi, C., Bateman, I. J., & Ferrini, S. (2014). Using revealed preferences to estimate the value of travel time to recreation sites. *Journal of Environmental Economics and Management*, 67(1), 58–70. <https://doi.org/10.1016/j.jeem.2013.10.003>
- 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. <https://doi.org/10.1016/j.tra.2013.12.010>
- Hensher, D. (2019). Using the Average Wage Rate to Assess the Merit of Value of Travel Time Savings: A Concern and Clarification. *Transport Findings*, 1–6. <https://doi.org/10.32866/5772>
- Hensher, D. A. (2001). The valuation of commuter travel time savings for car drivers: Evaluating alternative model specifications. *Transportation*, 28(2), 101–118. <https://doi.org/10.1023/A:1010302117979>
- Hensher, D. A. (2010). Hypothetical bias, choice experiments and willingness to pay. *Transportation Research Part B: Methodological*, 44(6), 735–752. <https://doi.org/10.1016/j.trb.2009.12.012>
- Hess, S., Choudhury, C. F., Bliemer, M. C. J., & Hibberd, D. (2020b). Modelling lane changing behaviour in approaches to roadworks: Contrasting and combining driving simulator data with stated choice data. *Transportation Research Part C: Emerging Technologies*, 112(May 2019), 282–294. <https://doi.org/10.1016/j.trc.2019.12.003>
- Hess, S., Daly, A., & Börjesson, M. (2020a). A Critical Appraisal of Use of Simple Time-Money Trade-Offs for Appraisal Value of Time Measures. *Transportation*, *forthcomin*.
- Hess, S., Daly, A., Dekker, T., Cabral, M. O., & Batley, R. (2017). A framework for capturing heterogeneity, heteroskedasticity, non-linearity, reference dependence and design artefacts in value of time research. *Transportation Research Part B: Methodological*, 96, 126–149. <https://doi.org/10.1016/j.trb.2016.11.002>
- Hess, S., & Palma, D. (2019). Apollo: a flexible, powerful and customisable freeware package for choice model estimation and application version 0.0.6 User manual. *Journal of Choice Modelling*. <https://doi.org/10.1016/j.jocm.2019.100170>
- Isacsson, G. (2007). *The trade off between time and money: Is there a difference between real and hypothetical choices?* (Working Papers 2007:3, Swedish National Road & Transport Research Institute).
- Knoop, V. L., Hoogendoorn, S. P., & van Zuylen, H. J. (2009). Processing Traffic Data Collected by

- Remote Sensing. *Transportation Research Record: Journal of the Transportation Research Board*, 2129(1), 55–61. <https://doi.org/10.3141/2129-07>
- Koutsopoulos, H. N., Lotan, T., & Yang, Q. (1994). A driving simulator and its application for modeling route choice in the presence of information. *Transportation Research Part C*, 2(2), 91–107. [https://doi.org/10.1016/0968-090X\(94\)90002-7](https://doi.org/10.1016/0968-090X(94)90002-7)
- Levinson, D., Harder, K., Bloomfield, J., & Winiarczyk, K. (2004). Weighting waiting - Evaluating perception of in-vehicle travel time under moving and stopped conditions. *Transportation Research Record: Journal of the Transportation Research Board*, 1898, 61–68. <https://doi.org/10.3141/1898-08>
- Louviere, J. J., & Hensher, D. A. (1983). Using Discrete Choice Models with Experimental Design Data to Forecast Consumer Demand for a Unique Cultural Event. *Journal of Consumer Research*, 10(3), 348. <https://doi.org/10.1086/208974>
- Meneguzzo, C. (2019). Contrarians do better: Testing participants' response to information in a simulated day-to-day route choice experiment. *Travel Behaviour and Society*, 15(February), 146–156. <https://doi.org/10.1016/j.tbs.2019.02.003>
- Papinski, D., Scott, D. M., & Doherty, S. T. (2009). Exploring the route choice decision-making process: A comparison of planned and observed routes obtained using person-based GPS. *Transportation Research Part F: Traffic Psychology and Behaviour*, 12(4), 347–358. <https://doi.org/10.1016/j.trf.2009.04.001>
- Peer, S. (2013). *The economics of trip scheduling, travel time variability and traffic information*. <http://trid.trb.org/view.aspx?id=1255531>
- Prato, C., Rasmussen, T., & Nielsen, O. (2014). Estimating Value of Congestion and of Reliability from Observation of Route Choice Behavior of Car Drivers. *Transportation Research Record: Journal of the Transportation Research Board*, 2412, 20–27. <https://doi.org/10.3141/2412-03>
- Qi, H., Ma, S., Jia, N., & Wang, G. (2019). Individual response modes to pre-trip information in congestible networks: laboratory experiment. *Transportmetrica A: Transport Science*, 15(2), 376–395. <https://doi.org/10.1080/23249935.2018.1485061>
- Ramos, G. M., Frejinger, E., & Hoogendoorn, S. (2012). A revealed preference study on route choices in a congested network with real-time information. *13th International Conference on Travel Behaviour Research, Toronto, Canada, 15-20 July, 2012*.
- Razo, M., & Gao, S. (2013). A rank-dependent expected utility model for strategic route choice with stated preference data. *Transportation Research Part C: Emerging Technologies*, 27, 117–130. <https://doi.org/10.1016/j.trc.2011.08.009>
- Rieser-Schüssler, N., & Axhausen, K. W. (2014). Self-tracing and reporting: State of the art in the capture of revealed behaviour. In *Handbook of Choice Modelling* (pp. 131–151). <https://doi.org/10.4337/9781781003152.00012>
- Rose, J. M., & Bliemer, M. C. J. (2009). Constructing Efficient Stated Choice Experimental Designs. *Transport Reviews*, 29(5), 587–617. <https://doi.org/10.1080/01441640902827623>
- Rosenthal, R. W. (1973). A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory*, 2(1), 65–67. <https://doi.org/10.1007/BF01737559>
- Saifuzzaman, M., Zheng, Z., Haque, M. M., & Washington, S. (2017). Understanding the mechanism of traffic hysteresis and traffic oscillations through the change in task difficulty level. *Transportation Research Part B: Methodological*, 105, 523–538.

<https://doi.org/10.1016/j.trb.2017.09.023>

- Senna, L. (1994). The influence of travel time variability on the value of time. *Transportation*, 21(2), 203–228. <https://doi.org/10.1007/BF01098793>
- Sharma, A., Zheng, Z., Bhaskar, A., & Haque, M. M. (2019). Modelling car-following behaviour of connected vehicles with a focus on driver compliance. *Transportation Research Part B: Methodological*, 126, 256–279. <https://doi.org/10.1016/j.trb.2019.06.008>
- Shires, J. D., & de Jong, G. C. (2009). An international meta-analysis of values of travel time savings. *Evaluation and Program Planning*, 32(4), 315–325. <https://doi.org/10.1016/j.evalprogplan.2009.06.010>
- Small, K. A. (1999). Valuation of travel-time savings and predictability in congested conditions for highway user-cost estimation. In *NCHRP Report* (Issue 431). National Academy Press.
- Smith, V. L. (1976). Experimental Economics: Induced Value Theory. *American Economic Review*, 66(2), 274–279.
- TfNSW. (2019). *Transport for NSW Economic Parameter Values* (Issue September).
- Tseng, Y.-Y., Verhoef, E., de Jong, G., Kouwenhoven, M., & van der Hoorn, T. (2009). A pilot study into the perception of unreliability of travel times using in-depth interviews. *Journal of Choice Modelling*, 2(1), 8–28. [https://doi.org/10.1016/S1755-5345\(13\)70002-2](https://doi.org/10.1016/S1755-5345(13)70002-2)
- Vacca, A., Prato, C. G., & Meloni, I. (2019). Should I stay or should I go? Investigating route switching behavior from revealed preferences data. *Transportation*, 46(1), 75–93. <https://doi.org/10.1007/s11116-017-9788-9>
- van Essen, M., Thomas, T., Chorus, C., & van Berkum, E. (2019). The effect of travel time information on day-to-day route choice behaviour: evidence from a real-world experiment. *Transportmetrica B: Transport Dynamics*, 7(1), 1719–1742. <https://doi.org/10.1080/21680566.2019.1699198>
- Varela, J. M. L., Börjesson, M., & Daly, A. (2018). Quantifying errors in travel time and cost by latent variables. *Transportation Research Part B: Methodological*, 117, 520–541. <https://doi.org/10.1016/j.trb.2018.09.010>
- Wardman, M., & Whelan, G. (2001). Valuation of improved railway rolling stock: A review of the literature and new evidence. *Transport Reviews*, 21(4), 415–447. <https://doi.org/10.1080/01441640010020115>
- Young, W., Sobhani, A., Lenné, M. G., & Sarvi, M. (2014). Simulation of safety: A review of the state of the art in road safety simulation modelling. *Accident Analysis & Prevention*, 66, 89–103. <https://doi.org/10.1016/j.aap.2014.01.008>