MEASURING THE EFFECTS OF DIFFERENT EXPERIMENTAL DESIGNS AND SURVEY ADMINISTRATION METHODS USING AN ATLANTA MANANGED LANES STATED PREFERENCE SURVEY

Submission date: 1 August 2007
Word count: 6792 words + 2 figures and 6 tables

Stephane Hess
Centre for Transport Studies
Imperial College London
SW7 2AZ London
United Kingdom
Telephone: +44(0)20-7594-6105
Fax: +44(0)20-7594-6102
stephane.hess@imperial.ac.uk

Corresponding Author:
Colin Smith
Resource Systems Group
55 Railroad Row
White River Junction, VT 05001
Telephone: (802) 295-4999
Fax: (802) 295-1006
csmith@rsginc.com

Stacey Falzarano
Resource Systems Group
55 Railroad Row
White River Junction, VT 05001
Telephone: (802) 295-4999
Fax: (802) 295-1006
sfalzarano@rsginc.com

Jevan Stubits
Resource Systems Group
55 Railroad Row
White River Junction, VT 05001
Telephone: (802) 295-4999
Fax: (802) 295-1006
jstubits@rsginc.com

ABSTRACT
With the vast majority of choice modelling studies making use of data collected using Stated Preference (SP) surveys, there is considerable interest in the benefit of advanced survey techniques, such as efficient designs. However, while there is a growing body of statistical evidence of the advantages of such advanced methods, there is still relatively little evidence from actual real-world studies. In this paper, we present the findings of a large study making use of SP surveys in the context of road tolling initiatives. The study has the rather special characteristic that the data collection made use of survey questionnaires using various different experimental design approaches. This allows us to directly compare the stated choice behaviour across designs. We observe significant differences in a number of crucial indicators, such as the value of time and elasticities with respect to income and trip distance, depending on what design is used. The analysis further shows that superior performance is obtained when making use of an efficient design, while also highlighting the disadvantages of random blocking approaches.

1. INTRODUCTION

The Georgia Department of Transportation (GDOT) is evaluating the addition of managed lanes to I-75 inside of the I-285 perimeter, I-20 to the west and east of I-75, I-85 to the northeast of I-75, and all of I-285, collectively known as "the Study Routes" and highlighted in Figure 1.

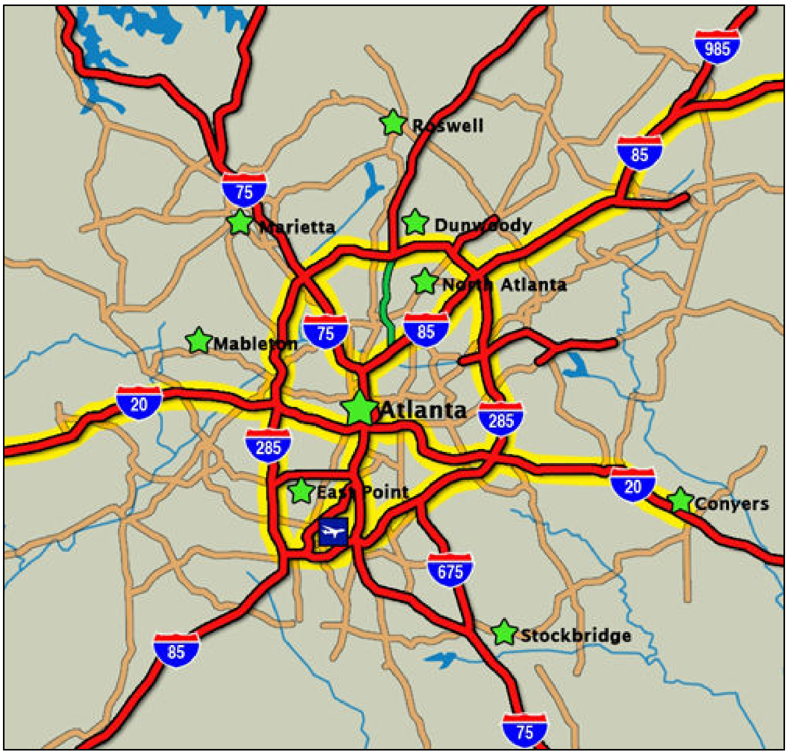

FIGURE 1 Map of the Atlanta region with the study routes highlighted

This paper describes stated preference survey design and data collection that was conducted for GDOT as part of the Managed Lanes System Study in May and June of 2007. The purpose of the stated preference survey was to obtain detailed information that could be used to determine how sensitive travellers would be to the tolling and travel-time changes that would result from managed alternatives on the Study Routes.

A stated preference survey was developed and implemented that gathered information from individuals who use the Study Routes. The survey collected data on current travel behaviour, presented respondents with information about the proposed managed lane alternatives, and, with the use of stated preference experiments, collected information that can be used to estimate travellers' valuation of travel time savings (VTTS) and propensity to use managed toll lanes or high occupancy vehicle (HOV) lanes under a range of possible future conditions.

Several experimental designs were tested during the stated preference survey. This paper describes the development and application of the experimental designs, which included an orthogonal design with random blocking, an orthogonal design with non-random blocking, and an efficient design. The paper presents a thorough analysis of the design outcomes, such as the time for respondents to complete the whole survey and to complete the stated preference experiments, non-trading within the stated preference experiments, and lexicographic behaviour (e.g. always choosing the fastest alternative). The paper also investigates the differences in model performance and results obtained using data collected with the different designs.

Along with the differences in survey design, it should also be noted that the survey was administered in several different locations. Here again, the paper describes differences in survey response and modelling results across these various locations.

Section 2 of this paper presents a detailed account of the data collection exercise. This is followed by a descriptive analysis of the data in Section 3, and by an analysis using discrete choice models in Section 4. Finally, Section 5 presents the conclusions of the paper.

2. STATED PREFERENCE SURVEY APPROACH

The stated preference survey was designed to identify the travel patterns and preferences of passenger vehicle travellers who could reasonably use a managed lane alternative in the greater Atlanta area.

The stated preference survey employed a computer-assisted self interview (CASI) approach. The stated preference survey instrument was customized for each respondent by presenting questions and modifying wording based on respondents' previous answers. These dynamic survey features provide an accurate and efficient means of data collection and allow presentation of realistic future conditions that correspond with the respondents' reported experiences.

In this section, we give a detailed account of the data collection exercise. We first present a description of the survey questionnaire. This is followed by a discussion of the experimental design techniques used to generate the choice situations for individual respondents. Finally, we turn our attention to survey administration.

2.1 Stated Preference Survey Questionnaire

Respondents were screened to ensure that they would describe trips that used the Study Routes in the greater Atlanta area. Respondents were asked if within the last week they had made a weekday trip of at least 15 minutes that used any of the Study Routes in the greater Atlanta area: specifically, I-20 from Villa Rica on the west to Conyers on the east, I-85 to Braselton to the north, I-285 and I-20, I-75, and I-85 within the I-285 perimeter. These screening criteria, in combination with validation of respondents' origins and destinations, ensured that respondents focused on a trip that could reasonably use the proposed managed lanes. Respondents were asked to keep this trip in mind as they completed the questionnaire.

The questionnaire consisted of four main parts: context questions that asked for details about the respondent's trip, a description of the proposed managed lane alternatives for the greater Atlanta area, stated preference questions that presented a managed lane or carpool alternative to their current route for making the trip in the future under a range of possible conditions, and debrief and demographic questions. We will now look at these subparts of the survey in turn.

2.1.1 Context Questions

The first section of the questionnaire acquired details from respondents about a recent trip that could use the proposed managed lane alternatives in the future. Revealed preference information such as roads used, vehicle type, trip purpose, day of week, time of day, total travel time, and trip frequency were obtained.

The questionnaire identified home- and work-based trips. Respondents indicate the locations where their trip began and ended using their choice of entering the street addresses or clicking on an area map. Approximate trip origin and destination locations were recorded in latitude and longitude. The shortest route that would use the proposed managed lanes of the Study Routes was calculated based on the origin and destination locations. The closest on and off ramps were then identified and the distance on the proposed managed lanes was calculated. Respondent input was validated to ensure that the trip could reasonably use the proposed managed.

Each origin and destination and each interchange on the managed lanes portion of the Study Route was associated with a specific Traffic Survey Zone (TSZ). Skim data were used to calculate zone-to-zone distances and travel times. Respondents' reported travel times were validated against the skim data and a warning was shown if their reported travel time was significantly longer or shorter than the expected travel time between zones based on the skim data calculations.

Respondents were asked to indicate whether they experienced delay due to heavy traffic during their trip and, if so, to identify the approximate delay time. The questionnaire asked respondents the number of occupants in their car, who was in the car, and why the respondent chose to carpool. The final context questions asked respondents if they used Georgia 400 (the only toll road in area) and if they currently own an electronic toll collection (ETC) transponder.

2.1.2 Stated Preference Questions

The survey provided a description of the proposed managed lane alternatives for the Study Route. Respondents were provided with information about how the proposed managed lanes would work, and they were told that toll free existing lanes would still be available.

The survey presented each respondent with eight stated preference choice situations each with two or three travel alternatives, depending on the current vehicle occupancy. Figure 2 shows an example choice situation presented to a traveller who described a trip where they drove alone. In this case, the three alternatives shown were labelled "Existing Lanes: Drive Alone," "New Managed Lanes: Drive Alone," and "New Managed Lanes: Carpool". Specific details of the alternatives were customized based on the reported travel time, toll cost, and vehicle occupancy for each respondent's current trip.

Each choice situation asked the respondent which of the three alternatives was most preferred for making the trip described. Each of the alternatives included information about the travel time, toll cost, and vehicle occupancy. Across all of their choice situations, the respondent was presented with different levels of each of these varied attributes (travel time, toll cost, and vehicle occupancy). The attributes and levels for peak period travellers are shown in Table 2; off-peak travellers were shown lower tolls and lower time saving levels in the managed lanes.

The specific values assigned in each choice situation were determined using an experimental design with 32 rows. A set of eight choice situations were selected from the design and presented to each respondent in random order. Several experimental designs and methods for selecting the eight choice situations for a particular respondent were tested during the survey; details of the experimental designs used are given in Section 2.2.

FIGURE 2 Screenshot of stated preference choice situation

TABLE 2 Stated Preference Attributes and Levels

Attribute	Existing Lanes Travel Time (Low and High Variation)[1]		Managed Lanes Travel Time[2]	Managed Lanes Toll[3]	Carpool Toll	Carpool Occupancy
Peak Period Level Values	- 7.5mph	- 15mph	+ 25mph	5c/mile	Toll Free	+1 Occ
	- 3.75mph	- 7.5mph	+ 30mph	10c/mile	ML Toll*0.33	+2 Occ
	+ 3.75mph	+ 7.5mph	+ 35 mph	15c/mile	ML Toll*0.67	
	+ 7.5mph	+ 15mph	+ 40mph	20c/mile	ML Toll	
				25c/mile		
				30c/mile		
				35c/mile		
				40c/mile		

1

2

3

The managed lane travel time was based on the respondent's reported travel time, with time savings proportional to the distance the respondent would travel on the proposed managed lane alternative on the Study Route. The carpool alternative's travel time was equal to the

[1] Speed varied around reported speed on Study Route; time on roads other than Study Route not varied. Lower variations used when reported speed 15mph or less. Higher variation used when reported speed 50mph or more. Variations interpolated for reported speeds between 15mph and 50mph.
[2] Speed varied based on reported speed on Study Route; time on other roads not varied.
[3] Toll based on mileage on Study Route.

managed lane travel time plus an additional three minutes per passenger in the carpool to allow for carpool formation. The existing lanes travel time was based on reported travel time, with time increases proportional to the distance the respondent would travel on the managed lanes.

The toll cost for the managed lanes alternative was proportional to the distance the respondent would travel on the managed lanes. The levels tested varied between 5 cents per mile and 40 cents per mile at peak travel times and between 2 cents per mile and 23 cents per mile at off peak travel times. The toll cost for the carpool alternative was a discount between 0% (i.e. the same as the managed lanes toll) and 100% (i.e. the carpool lane was toll free) of the managed lanes toll cost.

The vehicle occupancy tested in the carpool alternative was the addition of either one or two passengers to the current (reported) occupancy for the trip. The carpool alternative was not presented in the choice situations if the respondent was already travelling in a vehicle with three or more occupants.

With any stated preference survey, there is a risk of non-trading behaviour, i.e., respondents always choosing the same option, and this can have significant impact on model results. To alleviate these issues, respondents who always choose a non-toll option were presented with a ninth choice using the lowest presented value of time savings but with a reduction in the toll by 50%. Similarly, respondents who always chose the tolled alternative were presented with a ninth choice situation using the highest presented value of time savings with a doubling of the tolls.

2.1.3 Debrief Questions

At the conclusion of the stated preference questions, respondents were asked why they did or did not choose the managed lane options driving alone or as a carpool. Respondents were then asked for their opinions about the proposed managed lane scheme in the greater Atlanta area. The first question asked participants to indicate their overall support or opposition for the project. Respondents who said they "strongly favour" or "somewhat favour" the project were shown a follow-up question asking the primary reason they are in favour of the project. Respondents who said they "somewhat oppose" or "strongly oppose" the project were shown a follow-up question asking the primary reason they oppose the project.

2.1.4 Demographic Questions

In the last section of the survey, general demographic questions were asked to allow comparison of the sample population to the overall population that would be served by the proposed highway improvements. The demographic questions included resident/visitor status, county of residence, household size, number of household vehicles, gender, age, employment status, access to the Internet, point of Internet access, and annual household income.

2.2 Experimental Designs

A series of variations in the experimental design and the method for selecting the set of choice situations for a particular respondent were tested during the administration of the survey. Specifically, four main stages were used. These were:

- Stage 1: No variation in the existing lanes alternative across choice situations, with the managed lanes and carpool alternatives constructed using an orthogonal experimental design and random blocking to select the experiments for each respondent

Stage 2: Orthogonal experimental design and random blocking for all alternatives
Stage 3: Orthogonal experimental design and non-random blocking for all alternatives
Stage 4: Efficient experimental design with inherent balance, orthogonality, and non-random blocking

We will now look at the design approaches used in the four stages.

2.2.1 Orthogonal Design

The orthogonal designs used in this survey were constructed by selecting and combining interactions from a full factorial experimental design made up of five two level attributes. The 32 row experimental design contained one attribute with eight levels, three with four levels and one with two levels.

The first version of the orthogonal design (Stage 1) used no variation in the attributes for the existing lanes alternative across observations (for a given respondent). For all remaining respondents, independently of the design used (i.e., orthogonal or efficient), the attributes for the first alternative varied across observations.

However, a further distinction has to be made in the use of the orthogonal designs. This relates to the blocking used to select a set of eight choice situations out of a 32 row design matrix. For Stage 1 and Stage 2 of the survey, the blocking that was used was purely random, i.e. eight choice situations were selected randomly from the 32 available choice situations, where this random process was repeated for each respondent. In later experiments (Stage 3), the blocking was performed so as to minimize the correlation between attribute levels within a certain block, thus attempting to avoid the situation where one respondent obtains all the choice situations with low existing lanes travel times, while another respondent obtains all the choice situations with high existing lanes travel times. As discussed by Rose & Hensher (2006), this could lead to a loss of orthogonality.

2.2.2 Efficient Design

In recent years, the statistical state of the art of designing stated preference experiments has moved away from orthogonal designs to D-efficient designs (see for example, Burgess and Street, 2003; Carlsson and Martinsson, 2003; Ferrini and Scarpa 2006; Kessels *et al*., 2006; Sándor and Wedel, 2001; Street and Burgess, 2004)

A statistically efficient design is a design that minimizes the elements of the asymptotic (co)variance matrix, Ω, with the aim of producing greater *reliability* in the parameter estimates given a fixed number of choice observations. That is, a statistically efficient design is constructed with the aim of maximizing the asymptotic *t*-ratios obtained from data collected using the design. The construction of an efficient design therefore requires that the analyst construct the likely asymptotic (co)variance matrix for the design prior to collecting any data. This requires that the attribute levels and parameter estimates be known in advance.

Once the asymptotic (co)variance matrix has been constructed, it is necessary to determine the statistical efficiency of the design. To compare the statistical efficiency of SC experimental designs, a number of alternative approaches have been proposed within the literature (see e.g., Bunch *et al.* 1996). The most commonly used measure is the D-error.

$$\text{D-error} = \left(\det \Omega\right)^{1/k} = -\frac{1}{N}\left(\det\left(\frac{\partial LL(\beta)^2}{\partial\beta\partial\beta'}\right)\right)^{-1/k}. \qquad (1)$$

where k represents the number of parameters for the design, $LL(\beta)$ the log-likelihood function of the discrete choice model under consideration, N the sample size, and β the parameters to be estimated from the design.

A D-efficient design was generated for the purposes of the present study (Stage 4), using prior parameter values of –0.04 and –0.19 for the travel time and toll coefficients, giving a value of travel time savings of $12.60 per hour.

2.3 Survey Administration

Data collection was conducted in May and June 2007. A total of 4173 travellers who made a weekday trip of 15 minutes or more that used or could have used any of the major interstates within the greater Atlanta study area were intercepted at various activity locations or contacted via e-mail and completed the survey.

The computer-based survey was administered to four groups of respondents:

1. Respondents intercepted at activity locations in the greater Atlanta area using laptop-based administration.
2. Employees of large area businesses and public institutions using email invitations and Internet-based completion.
3. Patrons or visitors at activity locations in the greater Atlanta area, who received a postcard inviting them to take the survey via the Internet.
4. Survey panel members resident in the Atlanta area using email invitations and Internet-based completion

The survey questionnaire was administered at activity locations over a month long period between Thursday, 17 May, and Friday, 15 June with 1812 respondents completing the survey. The survey was administered at sites with high pedestrian traffic and high incidence of people likely to meet the screening criteria. Sites were chosen that would likely allow a good cross section of the population to be intercepted, in terms of both trip purposes and demographics. Survey sites included Georgia Department of Driver Services, shopping centres, universities, office buildings and sports events.

The intercept survey administration setup consisted of 20 laptop computer interview stations distributed across three to four locations each day. A poster mounted on an easel was positioned near the interview stations to help attract respondents. Each survey site was staffed by three attendants who were responsible for soliciting and screening potential respondents, escorting the respondents to interview stations, and assisting respondents with questions or use of the computers if necessary.

Employees from large corporate offices and public institutions located in Atlanta were invited by e-mail to complete the survey via the Internet. A link in the e-mail invitation brought them to a website hosted on RSG’s SurveyCafe.com. On the first page of the survey, respondents were provided with instructions for taking the survey, along with an email address and a toll-free telephone number to request assistance if necessary. A total of 1278 employees of local business and institutions completed the survey.

Some potential respondents intercepted at activity locations indicated that they did not have enough time to complete the survey at that particular time, but they were still interested in participating in the survey at a different time. They were handed a postcard with a link to the survey and a unique password. The link also brought them to the same website hosted on RSG's SurveyCafe.com. One-hundred seventeen respondents completed the survey via the postcard link.

Respondents were also recruited via e-mail from a survey panel. Survey panel members resident in the Atlanta area were invited to take the survey, and 966 completed the survey on the Internet

3. DESCRIPTIVE ANALYSIS

A descriptive analysis of the responses to the stated preference survey has been undertaken to investigate the effects of the experimental designs tested and differences in responses by survey administration method and location. In this analysis we review data such as the time taken by respondents to complete the whole survey and to complete the stated preference questions, non-trading within the stated preference questions, and lexicographic behaviour (e.g. always choosing the fastest alternative).

A total of 4173 responses were collected from a variety of sources[4]. The largest proportion (43%) came from respondents intercepted in the Atlanta area. Employees of large businesses, who took the survey via the Internet, comprised 24% of the sample. An additional 23% of respondents were recruited from a survey panel. Surveys were also completed by employees of large public institutions, and by individuals approached at field sites throughout Atlanta who later completed the survey via the Internet.

The 1812 individuals completing the survey at activity locations were intercepted at numerous sites throughout the greater Atlanta area. The majority (69%) was recruited at local malls and shopping centres and an additional 15% were intercepted at colleges and universities. The remaining surveys were completed at Departments of Drivers Services (12%) and office buildings (4%).

Nearly half of respondents (46%) completed the survey during stages 1 and 2, with an orthogonal experimental design and random blocking. Of these, 18% saw no variation in travel time for their current travel alternative while 82% were shown varying travel times across their set of choice situations. The remaining 54% of respondents were shown choice situations from an experimental design using non-random blocking – 51% of which saw experiments derived from an orthogonal design and 49% from an efficient design.

The experimental designs with random blocking were shown exclusively to respondents intercepted at activity locations in the Atlanta area. Nearly all of these respondents (95%) took the survey on a laptop at an activity location, while 5% received a postcard with a password and took the survey via the Internet at a later time.

The experimental designs with non-random blocking were shown to respondents completing the survey via the Internet. Respondents in this group were employees of area businesses and organizations or members of a survey panel. During internet administration, respondents were randomly assigned to be shown choice situations from the orthogonal or from the efficient experimental designs. Table 2 shows responses by design and source.

[4] For the present study, no segmentation by purpose was used. This also applies to the choice modelling analysis in Section 4. The development of more detailed, purpose-specific models is the subject of ongoing work.

TABLE 2 Responses by Design and Source

	Stage 1	Stage 2	Stage 3	Stage 4	Total
Activity locations	342	1470	0	0	1812
Postcard	12	93	2	10	117
Large businesses	0	0	497	515	1012
Public Institutions	0	0	101	98	199
Other business/institutions	0	0	35	32	67
Survey panel	0	0	511	455	966
Total	354	1563	1146	1110	4173

3.1 Random Blocked Choice Situations

Trip purposes were distributed evenly among respondents in the stage 1 and stage 2 (random blocked) sample. For respondents who saw variation in the existing lanes alternative (stage 2) as well as those who did not (stage 1), approximately 40% of trips were to or from work. The greatest difference was among school trips with more than 15% of non-varying sample trips going to/from school compared to 9% of the varying sample. At DDS and mall locations, more than half (60% and 56%, respectively) of trips described were to/from work or work-related business, compared to 36% among college/university respondents.

Almost half (47%) of the non-varying sample reported trips between 10 and 20 miles in distance compared to 33% of the varying sample. Less than half (42%) of respondents in the non-varying sample reported trips of 20 miles or more compared to 57% of the varying sample. More than half (52%) of university respondents reported a trip distance less than 20 miles compared to 45% of mall, and 40% of DDS respondents.

Just under half of non-varying and varying sample trips took place during the morning peak (7-10 AM) while respondents in the varying sample made 8% more trips than the non-varying sample during the evening peak (3-6 PM). Approximately half of mall, DDS, and university respondents reported an AM peak trip and approximately one-third reported off-peak trips.

Reported incomes were similar among the non-varying and varying random blocked samples. The largest difference was a 5% higher incidence of incomes below $25,000 for the non-varying sample. Nearly one-third (32%) of university respondents reported income under $25,000 while less than 20% of mall and DDS respondents reported this same income level.

Respondents in the non-varying sample spent a similar amount time taking the survey as those who were shown variation. Approximately 50% of both samples completed the survey in 7 to 13 minutes.

The descriptive analysis shows that relatively small demographic differences are present between the stage 1 and stage 2 sample, suggesting that any differences found between models estimated using the data from the two groups are unlikely to be due to sample composition. Larger demographic differences were observed between the samples from malls, DDS offices, and colleges/universities.

3.2 Non-Random Blocked Choice Situations

Similar distributions in trip purpose are observed among the orthogonal (stage 3) and efficient (stage 4) non-random blocked samples. Approximately 60% of both samples reported a trip to or from work, followed in size by social/recreational trips (14% in both samples) and other personal trips (10% in both samples). Just over 85% of large business employees and 77% of public institution employees reported a trip to or from work while only 29% of survey panel members reported the same trip purpose.

The orthogonal and efficient samples show similar distributions of trip distances, with 62% and 64%, respectively, reporting trips between 10 and 30 miles. However, differences exist among sources, with 74% of large business employees, and 85% of public institution employees describing a trip under 30 miles compared to 58% of survey panel members.

There is similarly little difference in orthogonal and efficient samples in terms of travel time of day and income. Meanwhile, the survey panel sample showed more than twice as many off-peak trips as the large business/public institution samples (29% to 12%) and large discrepancies exist among income, with 64% of large business employees, 26% of public institution employees, and 18% of survey panel members reporting incomes above $100,000.

Survey duration was distributed similarly in the orthogonal and efficient samples, with just under 40% of each sample completing the survey in less than 13 minutes. There were also only small differences in time taken to complete the stated preference questions between the orthogonal and efficient design samples, with 52% of each group completing the stated preference questions in 1 to 2 minutes. Total survey durations and stated preference questions durations varied by source. Nearly half (49%) of survey panel members completed the survey in less than 13 minutes compared to 31% for large business employees, and 25% for public institution employees. Similarly, 43% of survey panel members completed the stated preference questions in less than 90 seconds in contrast to 23% and 22% of large business employees and public institution employees, respectively.

As with the comparison between the stage 1 and stage 2 samples, the descriptive analysis shows that relatively small demographic differences are present between the stage 3 and stage 4 samples. Again, as demonstrated above, larger demographic differences were observed between the samples from large businesses, public institutions and the survey panel.

3.3 Comparison Between Random and Non-Random Blocked Choice Situations

Some demographic differences and differences in trip characteristics were observed between the sample that completed the survey during stage 1 and stage 2 (random blocking) at activity locations and the sample that completed the survey during stage 3 and stage 4 (non-random blocking) via the internet. These differences mean that we cannot fully discount sampling effects when comparing results for stages 1 and 2 with results for stage 3 and 4.

The random blocked sample contained fewer travel to work trips (40%) than the non-random blocked sample (59%). The time of day of travel of the random blocked sample was concentrated during the AM Peak (50%) and off-peak (31%), while the non-random blocked sample contained a higher proportion of PM Peak trips (47%). The travel time and trip distance distributions were similar in the two samples. The random blocked sample reported lower incomes (median income in the $50,000 to $75,000 range) than the non-random blocked sample (median income in the $75,000 to $100,000 range).

3.4 Non-trading and Lexicographic Behaviour

As a final stage in the descriptive analysis, we now look at the incidence of non-trading behaviour and apparent lexicographic behaviour. These two phenomena respectively relate to respondents always choosing the same alternative (e.g., existing lanes alternative) and respondents choosing on the basis of just a single attribute (e.g., always choosing the alternative with the lowest toll). The results of this are summarized in Table 3. Here, we present rates for respondents who always chose the same alternative across all eight choice situations, and respondents who always chose the fastest or cheapest option or the option with the lowest occupancy. Finally, we show a rate of additional non-trading for respondents who were presented with a ninth choice situation.

As a first observation, we can note that those respondents who consistently chose the same alternative across their initial choice situations maintained their non-trading behaviour into the ninth choice situation. Very few respondents always chose the fastest option, where this is correlated with over a quarter of respondents who always chose the cheapest option. Additionally, a large share of respondents never chose the carpooling alternative. In terms of influence of the design, we can note that the incidence of these various phenomena is increased when not allowing for variation in the attributes of the existing lanes alternative, where the rate of non-trading is also higher when using random blocking.

TABLE 3 Non-Trading and Lexicographical Behaviour by Experimental Design

	Non-trading over first 8 choices	Non-trading in 9th choice	Always choosing fastest option	Always choosing cheapest option	Always choosing option with lowest occupancy
Stage 1	44.35%	81.21%	4.03%	36.56%	62.10%
Stage 2	38.31%	78.58%	4.26%	25.78%	54.20%
Stage 3	31.73%	78.54%	1.39%	24.15%	55.73%
Stage 4	34.01%	74.50%	0.51%	26.70%	60.71%

4. CHOICE MODELING ANALYSIS

This section describes the findings of an analysis fitting discrete choice models to various subsets of the data collected for this study. With the aim being one of comparing results across different segments of the data, rather than fitting the best possible model structure to the data, simple Multinomial Logit (MNL) models were used. Nevertheless, as highlighted below, a relatively flexible utility specification was used, allowing for continuous interactions with income and trip distance information. All model results presented in this section were obtained using BIOGEME (Bierlaire, 2003).

4.1 Modelling methodology

For the present study, models were estimated in willingness to pay (WTP) space rather than preference space, so that the VTTS is estimated directly from the data, rather than having to be calculated on the basis of separate time and cost coefficients. This also has the advantage that a t-ratio is directly available for the VTTS.

With this in mind, the following base specification was used for the utility functions:

$$V_i = \delta_T I(T_i) + \beta_s \beta_V TT_i \left(\frac{inc}{\overline{\overline{inc}}}\right)^{\lambda_{VTTS,inc}} \left(\frac{inc}{\overline{\overline{inc}}}\right)^{\lambda_{scale,inc}} \left(\frac{dist}{\overline{\overline{dist}}}\right)^{\lambda_{VTTS,dist}} \left(\frac{dist}{\overline{\overline{dist}}}\right)^{\lambda_{scale,dist}} + \beta_s T_i \left(\frac{inc}{\overline{\overline{inc}}}\right)^{\lambda_{scale,inc}} \left(\frac{dist}{\overline{\overline{dist}}}\right)^{\lambda_{scale,dist}}$$

Various alternative-specific and individual-specific attributes enter this utility function. Specifically, we have that:

- $I(T_i)$ is set to 1 if alternative i has a non-zero toll
- TT_i gives the travel time of alternative i
- T_i gives the road toll of alternative i
- inc gives the income for the current respondent, with $\overline{inc}$ giving the base value
- dist gives the distance for the current respondent, with $\overline{dist}$ giving the base value

All remaining terms are estimated in the model. The term $\delta_{T,k}$ is a toll road penalty that can partly account for strategic bias. The term β_s is a scale coefficient, while β_V is the VTTS (in \$/min). The remaining terms are interaction terms, where $\lambda_{vtts,inc}$ gives the VTTS elasticity in relation to income, $\lambda_{scale,inc}$ gives the elasticity of the scale coefficient in relation to income, $\lambda_{vtts,dist}$ gives the elasticity of the VTTS to trip distance, and $\lambda_{scale,dist}$ gives the elasticity of the scale parameter to trip distance. The VTTS estimated in this model is for a trip length of $\overline{dist}$ and an income of $\overline{inc}$, with values of 26.37 miles and \$83261.9 dollars used in this analysis[5].

Some additional modifications of the utility functions are used in the estimation. An alternative specific constant (ASC) is included for general purpose (GP) lanes (alternative 1) and managed lanes (ML, alternative 2), while for car pooling lanes (CP, alternative 3), an additional coefficient is included when the occupancy rises by 2 passengers compared to the remaining two alternatives.

In some of the more advanced models, we allowed for differences in the relative weight of the observed and unobserved parts of the utility function across subsets of the estimation data. Here, the observed part of the utility functions was multiplied by segment-specific scale parameters, with the scale parameter in one segment being normalized to 1. A value for a scale parameter higher than 1 indicates that in the given segment, the variance of the unobserved part of utility is smaller than in the base segment.

4.2 Empirical analysis

Our empirical analysis is split into three main parts. In the first part, we compare the estimation results across the data subsets resulting from the different data collection stages. In the second part, we look at differences across survey locations. Finally, in the third part, we check for differences in scale across segments. We will now look at the different parts in more detail.

[5] For details on this elasticity specification, see Mackie et al (2003).

4.2.1 Differences across designs

In the first part of our analysis, 6 MNL models were estimated on the data. Specifically, the models were estimated on the following subsets of the estimation data:

- Full sample
- Stage 1: orthogonal design without variation for existing lanes alternative
- Stage 2: orthogonal design with existing lanes variation and random blocking
- Stage 3: orthogonal design with existing lanes variation and non-random blocking
- Stage 4: efficient design
- Stage 5: combining stage 3 and stage 4

The estimation results for the six models used in this section are summarized in Table 4. Here, some important differences arise across the various models. We can first observe that much better model fit is obtained with the stage 3 and stage 4 models, i.e. the models not using random blocking. The effects of the random blocking on model fit are very apparent from the stage 1 and especially the stage 2 models. The effects are slightly less extreme in the stage 1 models, which is quite probably caused by the fact that the attributes for the first alternative are constant across observations.

Another major difference arises when looking at the implied VTTS at the mean income and trip distance. Here, the results for the stage 3 and stage 4 models are very similar, while higher VTTS measures are obtained in the stage 2 model and especially the stage 1 model. Furthermore, while all models indicate that the VTTS increases with income and trip distance, there are again some differences across models in these elasticities. Here, we observe the most significant income effect in the models estimated on the stage 4 data (efficient design), where this is four times as large as the effect in the stage 1 and stage 3 models. The stage 2 and stage 3 models also indicate a much bigger effect of trip distance than the stage 1 and stage 4 models.

TABLE 4 Estimation Results for MNL Models for Different Experimental Designs

	ALL		STAGE 1		STAGE 2		STAGE 3		STAGE 4		STAGE 5	
Observations:	34,857		2,993		13,101		9,531		9,232		18,763	
adj. ρ^2:	0.3099		0.3112		0.2188		0.3876		0.3982		0.3920	
	est.	t-rat.	est.	t-rat.	est.	t-rat.	est.	t-rat.	est.	t-rat.	est.	t-rat.
ASC(GP)	1.9957	53.51	1.7374	13.17	1.6146	28.31	2.4186	30.90	2.3305	30.31	2.3773	43.38
ASC(ML)	1.3160	41.51	0.9039	8.72	1.1503	24.24	1.4817	22.78	1.5960	23.50	1.5398	32.78
δ_{toll}	-0.2301	-6.95	-0.5670	-5.25	-0.3809	-7.34	0.0205	0.31	-0.1040	-1.50	-0.0419	-0.88
λ_scale_dist	-0.9425	-26.82	-0.7612	-4.48	-1.0339	-13.77	-1.0615	-18.22	-0.9044	-16.04	-0.9824	-23.95
λ _scale_inc	0.0381	1.42	-0.3570	-2.50	-0.0530	-0.96	-0.1158	-2.43	-0.1301	-2.88	-0.1199	-3.65
λ _vtts_dist	0.2062	7.13	0.2073	1.80	0.3109	5.71	0.2829	5.58	0.1866	3.50	0.2434	6.68
λ _vtts_inc	0.1099	4.96	0.0904	1.00	0.2681	6.22	0.0844	2.26	0.3970	8.30	0.2194	7.52
$\beta_{occ,inc,2}$	-0.2104	-5.01	-0.1086	-0.78	-0.1313	-2.11	-0.1853	-2.15	-0.4034	-4.33	-0.2852	-4.52
β_S	-0.2880	-41.11	-0.2071	-5.93	-0.1808	-17.58	-0.4042	-25.96	-0.4054	-26.12	-0.4037	-36.87
β_V	0.1515	35.69	0.3272	6.15	0.2091	17.35	0.1255	22.95	0.1195	19.71	0.1231	30.47
VTTS ($/hr)	9.09		19.63		12.55		7.53		7.17		7.38	

In combination, these findings offer clear evidence of differences in results as a function of the experimental design used, with the most striking ones being the poor performance in terms of model fit for the stage 2 model, and the significant overestimation of the VTTS in the stage 1 model.

4.2.2 Differences across survey locations

In the second part of our analysis, we look at differences in estimation results across survey locations. Here, it is however important to cancel out as much as possible the differences caused by the fact that different designs were used more or less frequently for different locations. To this extent, we limited this part of our analysis to stage 5 observations, i.e. survey data collected either with the efficient design or the orthogonal design with non-random blocking. This leads to three subsegments with non-trivial sample sizes, namely the data collected:

- from large business employees;
- from survey panel members
- from public institution employees

The results for the three models estimated in this section are summarized in Table 5. Here, we can observe slightly poorer performance in the model for survey panel members, but the differences in fit across locations are in no way as dramatic as those observed across the different survey methods. The VTTS measures at the mean income and trip distance are also very comparable. However, we can see that the income effect is far greater for large business employees than for survey panel members, and even higher in the case of public institution employees, where for this final sample, the distance effect is less important.

TABLE 5 Estimation Results for MNL Models for Different Data Sources

	Large Businesses		Survey Panel		Public Institutions	
Observations:	8,402		8,032		1,662	
adj. ρ^2:	0.4206		0.3623		0.4390	
	est.	t-rat.	est.	t-rat.	est.	t-rat.
ASC(GP)	2.7649	33.10	1.9840	24.64	2.5755	13.14
ASC(ML)	1.8832	25.58	1.1508	16.91	1.6221	8.85
δ_{toll}	-0.0334	-0.49	-0.0538	-0.73	-0.0663	-0.33
λ _scale_dist	-1.1183	-16.32	-0.9346	-16.90	-1.0310	-6.66
λ _scale_inc	-0.1209	-1.93	0.0576	1.09	-0.2813	-1.96
λ _vtts_dist	0.2051	3.72	0.2414	5.06	0.1368	1.00
λ _vtts_inc	0.4445	8.05	0.0745	1.68	0.7321	4.51
$\beta_{occ,inc,2}$	-0.3871	-3.73	-0.2263	-2.55	-0.0327	-0.15
β_S	-0.3962	-22.79	-0.4602	-23.81	-0.3650	-9.02
β_V	0.1153	19.63	0.1278	20.20	0.1228	7.08
VTTS ($/hr)	6.92		7.67		7.37	

4.2.3 Differences in scale

In the final part of our analysis, we allow for differences in scale across the various segments. Here, two sub-analyses were conducted.

- First, we estimated a model on the full sample, allowing for differences in scale across the four stages of the data collection process, where stage 1, i.e. the survey without variation for the existing lanes alternative, was used as the base, with the scale parameter fixed to a value of 1.
- Second, we estimated a model only on the stage 5 data, where we also restricted ourselves to the three main survey locations for this stage, as listed above. Here, we allowed for differences in scale across the three locations as well as the two data collection stages in question (stage 3 and stage 4).

The results for the resulting two models are summarized in Table 6. In both cases, the inclusion of the scale parameters led to statistically significant gains in model performance[6]. The rescaling also has some minor effects on the mean VTTS, while it also leads to quite a major drop in the income elasticity in the model segmenting by survey location.

In the context of this section, the most interesting points can be made in relation to the scale parameters[7]. Here, we note that the weight of the unobserved part of utility is greatest for stage 3, i.e., the orthogonal design with non-random blocking, while it is lowest for stage 2 and stage 4, i.e. the orthogonal design with random blocking and the efficient design. This means that a greater weight for the observed part of utility does not necessarily translate into better model performance.

However, when additionally looking at differences in scale across survey locations, we can see that the difference between stage 3 and stage 4 is not constant across locations, where, for survey panel members, the scale is higher with stage 3 than with stage 4. The differences in scale across locations are however very stable, with the highest scale being obtained for large business employees, ahead of public institution employees and survey panel.

TABLE 6 Estimation Results for Models Allowing for Scale Differences

	By stage		By stage & loc.	
Observations:	34,857		18,096	
adj. ρ^2:	0.3185		0.3925	
	est.	asy. t-rat.	est.	asy. t-rat.
ASC(GP)	1.7475	31.30	2.5339	32.52
ASC(ML)	1.1608	28.07	1.6388	27.00
δ_{toll}	-0.1753	-6.01	-0.0585	-1.15
λ _scale_dist	-0.9811	-28.65	-0.9814	-22.69
λ _scale_inc	-0.0715	-2.77	-0.1940	-5.61
λ _vtts_dist	0.2408	8.31	0.2224	5.90
λ _vtts_inc	0.1962	8.79	0.2746	8.57
$\beta_{occ,inc,2}$	-0.1986	-5.24	-0.2914	-4.32
β_S	-0.2600	-27.96	-0.4264	-29.39

[6] Detailed comparisons available from the first author on request.
[7] For scale parameters, the t-ratios are calculate with respect to a value of 1, not 0.

β_V	0.1449	35.88	0.1244	29.85
Scale stage 1	1.0000	-	-	-
Scale stage 2	1.3846	8.70	-	-
Scale stage 3	0.8755	-4.49	-	-
Scale stage 4	1.3706	8.45	-	-
Scale Large Biz stage 3	-	-	1.0000	-
Scale Pub Inst. stage 3	-	-	0.9570	-0.83
Scale Surv Panel stage 3	-	-	0.8809	-3.88
Scale Large Biz stage 4	-	-	1.0828	2.32
Scale Pub Inst stage 4	-	-	0.9908	-0.17
Scale Surv Panel stage 4	-	-	0.7978	-7.25
VTTS ($/hr)		8.69		7.47

5. CONCLUSIONS

In this paper, we have presented evidence to show that the use of different survey design techniques can lead to significant differences in model results, and hence potentially affect policy decisions in different ways. Our study stands out in that the different design techniques were all used in the context of a single study.

The findings from our analysis show that the use of different design techniques not only leads to differences in model performance, but that there are also differences in the implied willingness to pay indicators, the elasticities thereof with respect to income and trip distance, and the relative size of the observed and unobserved part of utilities.

Specifically, our analysis has shown that the use of random blocking approaches leads to much poorer performance, along with an overestimation of the VTTS, where this is even worse when not using variation in the attributes (across observations) for the *reference* alternative. Either approach also increases the risk of non-trading behaviour. When using non-random blocking, the performance with the orthogonal design is close to that observed with the efficient design, and the overall results are quite comparable.

A point that can be made at this stage is that VTTS measures obtained from the models were slightly lower than expected, and also lower than the $12.60 prior VTTS value used as the priors in the efficient design. The difference between posterior and prior values should not be a cause for grave concern. Indeed, the prior values used at the design stage only serve as an indication of the true values, i.e. a step in the right direction. It should be remembered that using an orthogonal design is equivalent to making an assumption of zero priors, such that the estimated parameters are in this case clearly closer to the efficient priors than to the orthogonal priors.

The fact that, independently of the issue of prior values, the actual VTTS estimates were lower than expected can almost certainly be seen as a result of strategic voting, with a large share of respondents rejecting tolled alternatives despite the time savings. This would inevitably lead to an underestimation of the travel time coefficient and an overestimation of the toll coefficient, leading to an underestimation of the VTTS. To a degree, this bias is reduced through the inclusion of the toll dummy terms which capture some of the political voting.

It is difficult to generalise the findings from our analysis, and evidence from other studies is required, again making use of different design techniques within the same study. Here, some guidance can be offered for further research:

- In our study, we had some confounding of locations and designs, such that for example respondents intercepted in the field were always presented only with stage 1 or stage 2 questionnaires. In future studies, all combinations of survey design and location should be used.
- In our study, separate efficient designs were only generated for a reference peak trip and a reference off-peak trip, and respondent-specific questionnaires were generated on the basis of percentage variations. This deviates from an optimal approach of producing a respondent-specific design on the basis of a specific trip. This is likely to improve results but the computational cost of generating respondent-specific designs, along with the necessary waiting times, should be borne in mind.

ACKNOWLEDGEMENTS
The authors would like to thank John Rose for helpful suggestions in relation to the use of efficient design techniques.

REFERENCES

Bierlaire, M. (2003) BIOGEME: a free package for the estimation of discrete choice models, Proceedings of the 3rd Swiss Transport Research Conference, Monte Verità, Ascona.

Bunch, D.S., Louviere, J.J. and Anderson, D.A (1996) A Comparison of Experimental Design Strategies for Choice-Based Conjoint Analysis with Generic-Attribute Multinomial Logit Models, Working Paper, Graduate School of Management, University of California, Davis.

Burgess, L. and Street, D.J. (2003) Optimal Designs for 2k Choice Experiments. *Communications in Statistics, Theory and Methods*, 32 (11), 2185–2206.

Carlsson, F. and Martinsson, P. (2002) Design Techniques for Stated Preference Methods in Health Economics. *Health Economics*, 12, 281-294.

Ferrini, S. and Scarpa, R. (2006) Designs with a-priori information for nonmarket valuation with choice-experiments: a Monte Carlo study, Working Paper submitted to Journal of Environmental Economics and Management.

Kanninen, B.J. (2002) Optimal Design for Multinomial Choice Experiments, *Journal of Marketing Research*, 39, 214-217.

Kessels, R., Goos, P. and Vandebroek, M. (2006) A comparison of criteria to design efficient choice experiments, *forthcoming* in *Journal of Marketing Research*.

Mackie, P., M. Wardman, A. S. Fowkes, G. Whelan, J. Nellthorp and J. J. Bates (2003) Values of travel time savings in the UK, Report to Department of Transport, Institute for Transport studies, University of Leeds and John Bates Services, Leeds and Abingdon.

Rose, J.M. and Hensher, D.A. (2006) Accounting for individual specific non-availability of alternatives in respondent's choice sets in the construction of stated choice experiments, Stopher, P.R. and Stecher, C. (eds.) Survey Methods, Elsevier Science, Oxford

Sándor, Z., and Wedel, M. (2001) Designing Conjoint Choice Experiments Using Managers' Prior Beliefs, *Journal of Marketing Research*, 38, 430-444.

Street, D.J., and Burgess, L. (2004) Optimal and Near Optimal Pairs for the Estimation of Effects in 2-Level Choice Experiments, *Journal of Statistical Planning and Inference*, 118(1–2), 185–99.