PATREC

Planning and Transport Research Centre (PATREC)

FINAL REPORT

STATED PREFERENCE SURVEYS - EXPERIMENTAL DESIGN

Submitted to	WA DoP – Dr Renlong Han and Alan Kleidon
Prepared by	Brett Smith, Doina Olaru, Huang Ying, and Sharon Biermann
Submitted by	Doina Olaru
Date	3rd of February 2015
Version	FINAL

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This executive summary provides an overview of the context and the support for a number of recommendations for the Stated Preference (SP) survey associated with the upcoming PARTS 2016 data collection.

The primary purpose of the Stated Preference survey is to improve the mode choice component of Perth's strategic models. Revealed Preference (RP) attributes often exhibit little variability or are collinear and it is difficult to determine their contribution to choice. SP will provide more robust parameter estimates than can be uncovered using only RP data. The current STEM and ROM24 models – supported by RP data only – report insignificant parameter estimates (gaps in the model) or parameters that are lower in absolute value than the prevailing evidence (Taplin *et al*. 2014, p. 48). Additionally, the models are non-responsive to behavioural and policy changes (e.g., travel time variability, public transport crowding, and uncertainty of finding a parking bay at a PnR station or at the trip destination, *ibid*.). SP is the appropriate tool to address the gaps in the strategic models, but it needs to be integrated with RP.

This Executive Summary focuses on the practical implications of SP and does not provide an abridged version of the theoretical components of SP experiments and their design (see Section 2 – Literature Review – and Section 3 – Experimental Designs). Detailed recommendations on the use of SP are made in Section 4 – Stated Preference Survey for Perth, Section 5 – Survey Administration and Associated Costs and in the conclusion of the present report. These recommendations can be summarised in the following 13 points:

1. SP is to form part of the PARTS 2016 household travel survey. A base design, developed with the purpose of providing parameter estimates for the mode choice models in Perth's strategic models, and a policy design including attributes that are not currently incorporated in the models is recommended. (***Recommendation 1 Section 4.2***)

2. SP to be primarily applied to the journey to work or education travel purpose, given the stress placed on the transport infrastructure of Perth by commuting. (***Recommendation 2 Section 4.2***)

3. SP instrument should be designed to provide insight to support decisions in the identified key policy areas. The two policies that are suggested by the Steering Committee and further explored in this report are:

   - **Managing Congestion during Peak Hour -** Managing travel demand on limited road supply during peak hour including the use of financial instruments;
   - **The demand for a "new" public transport mode** – Forecasting the patronage for a new light rail service.

   The experiments should be extensions to a base design aimed at achieving ***Recommendation 1.*** **(*****Recommendation 3 Section 4.2*****)**

4. Pilot studies should be undertaken on the survey instruments to judge the quality and consistency of the designs, given their reliance on prior parameter estimates. **(*****Recommendation 4 Section 4.2*****)**

5. Taking into account the priority of ***Recommendation 1***, the SP component of the survey needs to be solidly embedded into the main survey instrument and the mode choice components within

Perth's strategic models should aim to make the most of data enrichment (Sections 2.2-2.3) to provide robust parameters. **(*Recommendation 5 Section 4.3*)**

6. SP choice sets and attribute levels should be conditioned on RP data using appropriate segmentation that takes into account the respondent's context. This is referred to as the **hybrid approach** to pivoting and conditioning, where the respondent's current alternative (RP) appears in each choice set, but with attributes conditioned by average levels for commuters with similar trip characteristics. **(*Recommendation 6 Section 4.3*)**

7. PARTS 2016 should make use of attitudinal scales to improve the estimation of choice parameters and to inform behaviour modification policy. The level of ***environmental concern, affective factors (the importance of the car complimenting one's self-image), compliance to social norms and willingness to adopt new technologies*** not only help identify the sources of preference heterogeneity in the choice data, but also give policy makers some handle on what message to send as part of behavioural modification projects. **(*Recommendation 7 Section 4.4*)**

8. SP should account for departure time and incorporate it in the data collection as a blocking factor. To transition towards a tour-based model, SP should record the afternoon constraints/activities that may affect mode choice for morning commute. This is considered a solution balancing the complexity of design and cognitive burden for respondent and the sample size**s** required (thus data collection costs). **(*Recommendation 8 Section 4.3*)**

9. SP choice sets provided in the report (Appendix 3) may be used for the PARTS 2016 RP-SP survey in their current form. If adjustments are required, attributes with similar ranges and prior parameters may substitute current attributes. **(*Recommendation 9 Section 4.4*)**

10. To aid the consistency of communication with the respondents, the SP instrument should be as seamless as is possible, in that it looks and feels part of the main survey. The exact method for administering the SP instrument depends on the choice of instrument made by the survey agency administering PARTS 2016. However, a computer-based tool should be used. **(*Recommendation 10 Section 5.2*)**

11. A RP-SP PARTS 2016 is a state-of-the-art data collection exercise of large scale[1]. The administration of the surveys should be integrated into one organisation. WA DOT and DOP need to seek further information on the best organisational structure to meet the requirements of this joint data collection exercise **(*Recommendation 11 Section 5.3*).** Part of the tender process may ask potential contractors to suggest ways that they will meet the capability requirements.

The integration is recommended for the following reasons:

- **Organisational boundaries**: The design of SP relies heavily on access to RP data in a timely fashion. Cross-organisation boundaries will inhibit the rapid and efficient sharing of data.
- **Cost of administration:** An integrated RP and SP survey undertaken by one agency has the potential to reduce the cost of administrating both survey components. Given that the data collection agency would need to follow up with households to verify data in the travel diary or from the GPS records, the potential saving comes from incorporating the SP data collection into the validation process.

[1] Comparable SP-RP surveys in terms of depth and complexity are rare, with very few national examples in Europe (e.g., The Netherlands).

- **Continuity of conversation with the respondent:** The respondents will view the whole exercise more favourable if there is a consistent channel of communication; having a single organisation (branding, contact addresses, newsletters, and hearing complaints) is less frustrating to the respondent and will improve the quality of the data.

12. A business process for extracting, cleansing, formatting, modelling and design to be created to make sure the SP component is ready to go when validation of RP takes place. This process will depend on the choice of organisational structure as outlined above **(*Recommendation 12 Section 5.2*)**

13. The large scale of the undertaking along with the possibility that the RP survey may be conducted using new technology suggests a degree of risk. It is important that DOT/DOP appoint a client-side project management team that oversees the benefit realisation **(*Recommendation 13 Section 6*)**

1 INTRODUCTION

1.1 BACKGROUND AND AIM

This proposal is in response to the call of the Department of Planning, Western Australia (DoP), to design a suite of Stated Preference (SP) surveys and to examine the advantages and feasibility of combining them with the upcoming Metropolitan Perth and Peel Household Travel Survey (PARTS), scheduled to commence 2016. The aim of the SP surveys is to provide adequate data for estimating parameters for the two strategic transport models, currently used in our state (STEM and ROM24). The research project offers a framework for the evaluation of various survey approaches and suggests possibilities for collecting data necessary for making the components of STEM, ROM24 or the newly proposed PLATINUM strategic transport model more responsive to the current needs. This follows the recommendation from the review of the strategic model (Taplin *et al*. 2014).

The work undertaken by PATREC covered the following:

- An update on the literature review regarding the design, implementation and use of stated preference surveys;
- A number of designs for several parameter estimates that need updating or calculation in the strategic models. The analysis includes parking, a new transport mode and a financial measure such as congestion charging;
- Suggestions on data collection aspects, such as sampling, costs, protocol, etc.

1.2 STRUCTURE OF THE REPORT

The report has the following structure: Section 2 provides a literature review on Revealed Preference (RP) and Stated Preference (SP) surveys and provides a framework for their combination. Section 3 discusses the types of experimental designs and the challenges associated with them. Section 4 provides several considerations for designing experiments and describes three templates for SP surveys that could be undertaken at the same time with PARTS. Section 5 addresses the sampling issues and costs associated with SP surveys.

2 LITERATURE REVIEW

2.1 REVEALED AND STATED PREFERENCE DATA

The type of data collected in travel surveys is dictated by the type of question that the modeller attempts to answer. Two main alternative types of data are available to study travel behaviour: revealed preference and stated preference, with distinct benefits and limitations. Although SP data is not about observing what people actually do, the SP surveys can offer *"data consistent with economic theory, from which econometric models can be estimated, which are indistinguishable from their RP data counterparts"* (Louviere *et al*. 2000, p.21).

In transport modelling, RP data refers primarily to *ex post* information, i.e. trips or actions already undertaken by individuals, whereas SP is based on *ex ante* information, as the choice under investigation (and for which the subjects indicate the potential response) may occur in the future.
This is an important distinction that needs to be made, as the benefit of SP lies primarily on exploring new responses in hypothetical situations (new services, new markets), whereas RP's strength is in describing real state of affairs or options). Although RP has strong face validity (it describes the real state of affairs), it has numerous limitations:

- RP cannot be used to examine the transport impact of new alternatives;
- Lack of reliability when there is insufficient variability in the attributes;
- Highly collinear attributes make it difficult to determine the part worth of each attribute;
- Other noise around the elements that may have contributed to the selected course of action.

In addition, researchers have shown the importance of perceptions, preferences and attitudes in making decisions (Ben-Akiva *et al*. 2002), hence SP surveys would provide the solution for testing their relevance in Perth and embedding these behavioural elements into the PLATINUM strategic model proposed for WA. Whether the welfare analysis - using perceptions and attitudes data - will be significantly superior to the analysis based only on objective measures, this is a question that needs to be tested in the local context of WA.

2.1.1 Differences between RP and SP surveys

The main features of the two approaches are outlined in Table 1.

Table 1: Revealed and Stated Preference Features

Revealed Preference (RP)	Stated Preference (SP)
Respondents are required to answer a question about an action they performed; a problem with this approach is the reliance on the accuracy of information provided by the respondent (recall, post-rationalisation, perception of attributes of the alternatives).	Respondents are required to evaluate alternatives on a scenario and express their preference for one alternative[2]; this approach relies on experimental conditions manipulated by the researcher and presented to the respondent in a scenario.
Efficient when explanatory variables have substantial variability in the marketplace and also have little collinearity.	Useful for new attributes, services, hypothetical situations.
The choice of alternative is clear (what the individuals actually did), however inferring the relevant attributes and their values, and the potential options the respondents may have had is	The representation of the choice situation eliminates the noise around the attribute levels (an array of attributes can be considered), but the "actual choice" may not be what the respondents indicated as best for

[2] More recently, Best-Worst methodologies elicit from respondents not only the most preferred alternative, but also the least, and even the second best and second least (Marley 2009; Marley and Islam 2012; Rose 2013; Louviere *et al*. 2013).

Revealed Preference (RP)	Stated Preference (SP)
problematic; data is expected to have substantial amounts of measurement error.	them.
Subjects are asked only about their action and possibly about other options they may have considered.	Subjects are questioned about a sample of experiments/packages/bundles of features from the universe of possible experiments.
The attribute values may take any values and sometimes the limited variability of attributes may deem them insignificant in the models.	The attribute values are varied in a systematic fashion.
Embodies constraints of the decision maker; RP has high 'face validity'.	It cannot easily (if at all) represent changes in personal constraints and in the market.
As any other survey, the RP instrument needs testing, but not as extensive as the SP scenarios.	It involves a substantial preparatory phase between the design of experiments and survey (test, piloting, redesign of instrument).
One record per respondent.	Multiple observations are typically obtained from a respondent.
Limited statistical significance of certain parameters in cross-sections.	Generally more significant parameter estimates, as a result of the design manipulations (smaller standard errors).

The current direction in modelling is to combine sources of data, which was successfully applied in many research projects (Louviere *et al*. 2000 refer to this as 'data enrichment'). The potential modelling difficulty is in estimating the scale parameter for the SP and RP data sets, but substantial scholarly work provides guidance on how to achieve it.

2.2 COMBINING SP AND RP

2.2.1 Data enrichment

Basic Data Enrichment

Basic data enrichment techniques forecast travel behaviour responses based on RP data. SP data is used to improve the parameter estimates for attributes where RP provides insufficient variation or exhibits some other problem as outlined in Table 1. Morikawa (1989) and Ben-Akiva and Morikawa (1990) pioneered the enrichment methodology by assuming the taste parameters are the same for both RP and SP choice models, but the scale parameter may differ. The scale parameter relates to the variance of the unobserved effects (error) and therefore is not expected to be equal for each data set. A likelihood test using between two separate MNL models – one each for RP and SP – and a scaled adjusted pooled data MNL can be used to test the Morikawa assumption of equal parameters in each data set (Swait and Louviere 1993). In some cases, only some of the parameters are consistent across data sets, thus the calibration may include some generic and some data source-specific parameters. This is known as partial data enrichment.

The basic enrichment model estimates two MNL choice models that have common parameters, but a separate scale parameter accounts for the different variances in the unobserved utility. However, variations are possible. Yanez *et al.* (2010) modelled SP-RP models but considered the possibility that the underlying structure for each data set was nested. Hensher (1998) allowed the scale parameter for each alternative within each data set to be estimated freely by way of the heteroscedastic extreme value (HEV) choice model. In summary, the base model for each data set can be as flexible as the needs be, but a scale parameter adjustment is the way in which the data pooled into a single modelling exercise.

SP parameters and RP market shares: An alternate method of calibration is to estimate taste parameters using the SP data only (Swait *et al*. 1994). It does so by first estimating the parameters using the SP data. The estimation of market shares make use of RP data, but the taste parameters are fixed and alternative specific constants are estimated (i.e. recalibrated to reproduce market shares) constraining the estimates of taste parameters.

The basic data enrichment techniques treat each SP observation as being independent (no panel data effect).

Basic Data Enrichment Key Points

- Forecasts are based on RP data set, SP data is used to modify the parameter estimates;
- A RP choice observation, including the measurement of the attributes of the non-chosen alternatives needs to be recorded;
- The RP and SP should have a substantial overlap of the attributes;
- SP observations do not need to have similar attribute levels to RP data, but it makes sense to calibrate the attribute levels to RP;
- At least one SP choice observation is required. However it is more common to use more than one SP observation per respondent;
- If there are many SP observations and only one RP observation, some weighting function needs to be used. The weights usually are applied to give RP and SP equal importance.

Advanced Data Enrichment

Advanced data enrichment takes into account the panel data effects (i.e. all SP observations from one person are in some way related). The basis of the technique is to allow the scale parameter to Brownstone *et al.* (2000) estimate a RP-SP choice model using mixed-logit structure, which accounted for error correlation due to "repeated choices", preference heteroscedasticity, and scaling difference. Two estimation methods were proposed. Firstly, estimate the SP scale parameter using a basic enrichment technique and pre-multiply the SP variables by the estimated scale parameter. Secondly, estimate the preference parameters and the scale parameter in a full information maximum likelihood (FIML) model.

Bhat and Castelar (2002) extended the joint RP and SP enrichment modes by:

- Accommodating inter-alternative error structure (e.g. nested logit type models for each RP and SP sample);
- Accounting for the scale parameter difference between data sets;
- Estimating random parameters to capture preference heterogeneity;
- Accounting for state dependence (the tendency to choose the current RP mode among the SP alternatives) as well as heterogeneity in the state dependence.

The use of mixed logit models (either random parameter, RPL or error components, ECM) represented a significant advance in combining RP and SP data because the choices made by each individual are treated as unique to that individual, i.e. correlation between the unobserved components of the utility is accommodated. Hensher (2008) and Hensher *et al*. (2008) compared the results from an error components logit model (advanced data enrichment) and the nested logit technique (basic data enrichment) and concluded that advanced models explain the choices to a greater degree than basic enrichment models.

Advanced Data Enrichment Additional Points

- Advanced data enrichment models inherit the first five key points from basic data enrichment;
- Weighting the importance of RP and SP is not mentioned in the literature. However, it still may play a role and should be investigated;

- **(Main point)** The choice data for each individual is modelled using a panel data choice model. Any specification of choice heterogeneity is consistent across all SP and RP observations;
- The correlations between alternatives are modelled using an error component rather than nested logit.

2.2.2 Using RP to Grounding the SP Attribute Levels

Pivoting SP attribute levels on the RP observation

Individuals use availability and familiarity in their decision-making processes, thus it is considered more appropriate to design surveys that are close to respondent's actual choices (DeShazo *et al*. 2009). Pivoting the SP attribute levels means that the SP levels are proportions of the current RP choice, i.e., +25% or -25%. Pivoting can be around a reference alternative (the levels of the current choice by the respondent remain fixed) or an experienced alternative (the levels of the current choice by the respondent form part of the experimental design):

A reference alternative – the respondent's current RP choice may be used. The reference alternative *maintains the same attribute levels for every scenario*. The choice task becomes one of 'would you switch?' The reference alternative is somewhat limited in straight mode choice studies because policies that affect the respondent's mode are also of interest. However, when exploring the uptake of a new alternative (i.e., toll road or light rail) the reference alternative approach is appropriate (Hensher 2010).

An Experienced Alternative[3] - A less strict form of referencing is to pivot all SP attribute levels around an experienced alternative. This only differs from the referenced alternative approach because the attribute levels of the experienced alternative are included in the design. Basing the range of attribute levels around the respondent's experience reduces cognitive biases and improves design efficiency (Rose *et al*. 2008).

Conditioning SP attribute levels on the RP observation

Pivot designs present attributes that are centred on the attribute levels for each individual. It is also possible to condition the attribute levels on a reasonable expectation of what the attribute levels may be for an individual. Conditioning is appropriate when RP observations are not collected as part of the survey, or that the RP data is collected for one sample and the SP data is collected for another (in transport studies it is expected that both samples are in the same study area).

- **RP data not collected:** If there is no related RP data collection, conditioning SP attributes would be based on network data. This a similar approach to the way in which the attribute levels are specified for non-chosen RP alternatives ("imputation"). The analyst must decide the degree of disaggregation. A set of network attributes could be calculated for each household identified in the SP survey or some geographical boundary may be used, such as the suburb.
- **RP data collected for different sample:** The attribute levels for the SP sample are conditioned on averages from the preliminary RP data collection. Should the SP survey be separate from PARTS 2016, it is recommended that PARTS is completed first and any follow up survey is conditioned on the RP travel data. ***Note: this is not data enrichment, but a way to improve the face validity and forecasting capacity of a stand-alone SP survey.***

---

[3] The literature does not make the distinction between a referenced alternative (fixed attribute levels) and an experienced alternative (attribute levels are part of the experimental design). It should.

Conditioning the SP choice sets and attributes levels – a hybrid approach to pivoting and conditioning.

The LP110201150 - Modelling and evaluating the joint access mode and train station choice (PnR) used a hybrid of experiential alternatives. The respondent's current alternative (car, public transport alone and park-and-ride) appeared in each choice set, but with attributes conditioned by the average levels for commuters with similar trip lengths (short, medium and long). In this way, subjects responded to choice sets that included their current alternatives, but with attribute levels centred on an average rather than their own experienced levels. The strategy required nine experiments to be designed, being far less than one design for each respondent, but more realistic than ignoring the respondents current choice and context.

Pivoting in real time

The RP choice must be recorded before pivoting on a **reference or experienced alternative.** A mail-out survey will need two rounds of data collection. In between the collections the analyst needs to:

- complete the non-chosen alternatives data by using network information to obtain attribute levels,
- estimate RP choice models to provide prior parameter values for the SP design, and
- to pivot the attribute levels for the non-chosen alternatives[4].

The experience with the LP100100436 - Analysis and modelling of driving patterns for limited range electric vehicles, EV and the LP110201150 - Modelling and evaluating the joint access mode and train station choice (PnR) suggests **at least 12 weeks will elapse between survey rounds**.

Computer aided survey instruments have the potential to serve pivoted SP scenarios in real time. However, the programming task is non-trivial. The SP alternatives will select from a database of 'experimental designs', customised to respondents' circumstances. A matrix of designs with blocking factors such as purpose or distance could be used.

2.3 IMPLEMENTING STATED CHOICE EXPERIMENTS - HENSHER

Professor Hensher makes the following recommendations *"derived from …empirical evidence, carefully argued theoretical and behavioural positions, and speculative explanation"* (Hensher 2010, p.747)[5] :

1. The inclusion of a well-scripted presentation (including cheap-talk scripts), explaining the objectives of the choice experiment.
2. Inclusion of the opt-out or null alternative, avoiding a forced choice setting unless an opt-out is not sensible.
3. Pivoting the attribute levels of a choice experiment around an alternative that is the respondent's current choice or conditioning attributes on and alternative in which the respondent has a substantial awareness of.
4. Estimating unique parameter estimates for the reference alternative, in order to calculate estimates of marginal willingness to pay for an alternative that is actually chosen in a real market.
5. The ability to calibrate the alternative-specific constants through choice-based weights on alternatives where actual shares are known.

[4] If an experienced alternative is used, then the design will also include attribute levels for the current RP choice.
[5] The points are slightly reworded but text remains substantially the same as the original text by Hensher.

6. The inclusion of supplementary questions designed to identify the attribute processing strategy adopted, as well as a question to establish *“the confidence with which an individual would hypothetically purchase or use the good (or alternative) that is actually chosen in the choice experiment”*.

Data enrichment techniques aim to provide better forecasts of actual behaviour. The role of SP is to tease out the individual’s preference functions or decision rules. The recommendations made by Hensher (2010) aim to bed the SP task within the real context. This aim should be pursued in an SP exercise aimed to augment a household trip diary survey.

2.4 FORECASTING USING RP-SP MODELS

The golden rule of forecasting with RP/SP enrichment models is that the forecast is to be based on revealed preference shares and attribute levels. Enriched parameter estimates (i.e. whereby the same parameter is estimated from both SP and RP data, taking into account scale differences) are preferred, but if not possible, SP parameters conditioned on RP should be used in forecasts (based on RP data set).

The usual forecasting exercise is to examine what are the behavioural responses to a travel demand or supply policy instrument. The basis of such a forecast must be the current conditions, which is found in the RP data. However, three issues may arise in RP/SP collection and modelling that mean that the forecast cannot strictly be based on RP:

1. A new alternative or untested policy is introduced into the experiment. The market shares and any alternative specific parameter related to the new alternative are unknown. The alternative specific constant estimated on the SP data is rescaled (multiplied by the scale parameter) and the alternative is added to the RP data.
2. Should one or more of the RP and SP parameters to be significantly different, a partial enrichment model is specified. This leaves the decision maker with a choice on which parameter to use for forecasting. Conventional wisdom says that the RP parameter estimates take precedence.
3. The systematic or random heterogeneity structure of the RP and SP components of the model differs. The choice forecasting model is largely a decision to be made by the analyst. Cherchi and Ortúzar (2011) suggest testing the integrity of the forecasting estimates. For example, the marginal utility of travel time is expected to be negative. The choice on whether the RP or SP attribute levels may be guided by choosing the one that minimises the percentage of respondents for which the computed marginal utility is positive.

Taking into account the guidelines for implementing SP surveys (Hensher 2010; listed in section 2.2.6) and the forecasting guidelines (Cherchi and Ortúzar 2011; listed above) it is recommended that the SP component of the Perth survey integrates with PARTS 2016. The revealed preference results will form the basis of the design as well as producing models based on combined RP and SP data. This recommendation is formalised in **Section 4.2**.

2.5 COLLECTING INFORMATION ON ATTITUDES

In addition to the large body of work on properties of measurement scales (Louviere *et al*. 2000; Rose 2013), it has been shown that SP surveys seek to obtain highly disaggregated data on the decision maker characteristics and their attitudes and preferences. This allows for estimating hybrid choice models (Bolduc and Alvarez-Daziano 2013), relevant for market appraisals, as they account for various patterns of decision makers facing similar choice conditions.

The attitudinal survey instrument will depend on the context of the inquiry. For example, attitudes towards the environment may be relevant for the introduction of a new public transport mode, whereas attitudes to risk and uncertainty may be explored when examining congestion management instruments. A formal recommendation to include attitudinal measurements in the SP component of the household travel survey is made in Section 4.4 and examples of possible items that could be incorporated in the PARTS 2016 are offered in Appendix B. These questions are expected to better explain individual behaviour and they address social influences or subjective norms, perceived usefulness of new technologies, concerns about unsustainable use of resources and climate change, or lifestyles. Attitudinal constructs, expressed as latent variables, could be tested as covariates inside the utility functions for the discrete choice models.

3 EXPERIMENTAL DESIGNS

Choosing is a *"ubiquitous state of activity in all societies"* (Louviere et al. 2000, p.1) and individuals choose between different things based on their preferences/liking, personal situations, and experiences, accounting for their constraints and habits or inertia. Discrete choice experiments are used to elicit respondents' choices between a finite set of alternatives based on their preferences (Amaya-Amaya *et al*. 2007; Ortúzar and Willumsen 2011). The surveys may be delivered using different means, such as pencil and paper, computer aided personal interviews, or Internet. In a choice experiment, the attribute combinations attached to labelled or unlabelled alternatives are presented to the respondents in choice scenarios. Some experiments include a reference alternative in which the attributes of the experiment design are relative to the attributes of the respondent's actual choice in the market.

The concept of experimental design is at the core of SP choice surveys. Similar to physical sciences lab experiments, the SP experiments make possible fixing circumstances except for the issue in question. A good design is one that has a sufficiently rich set of attributes and choice contexts, together with enough variation in the attribute levels in order to produce meaningful behavioural responses. These conditions stem from the definition of experiment as a *"way of manipulating attributes and their levels to permit rigorous testing of certain hypotheses of interest"* (Louviere *et al*. 2000, p.84). The design seeks to address four objectives: 1) identification (the form of the utility function that can be estimated from a given experiment); 2) precision (smaller confidence intervals on parameter estimates); 3) cognitive complexity (managing task complexity); and 4) marketing realism (how close the experiment resembles the market conditions).

Briefly, an experimental design starts with identifying the relevant characteristics/attributes of the alternatives that will be compared in the scenarios. Then the number of levels and ranges of variables are set for each alternative, keeping in mind a relative balance (for example, attributes with a higher number of levels draw the attention of the respondent to those attributes, which may then result in a non-compensatory behaviour), the possibility of dominance (a bundle that is always better than the others), or that some combinations of attribute levels may be implausible (e.g., low fuel consumption and large engine size) or the combinations of levels give rise to interactions. Next, combinations/bundles of attributes are selected for each alternative and mixed into scenarios (also called stimuli or games) that will be shown to the respondents. The manner in which these combinations are built is the focus of the design of experiments.

3.1 STEPS REQUIRED IN SETTING UP SP DESIGNS

Figure 1 presents the sequence of activities required to be undertaken in experimental design for discrete choice models. Although the diagram does not specifically address the issue of optimisation, it provides a framework for building any SP design.

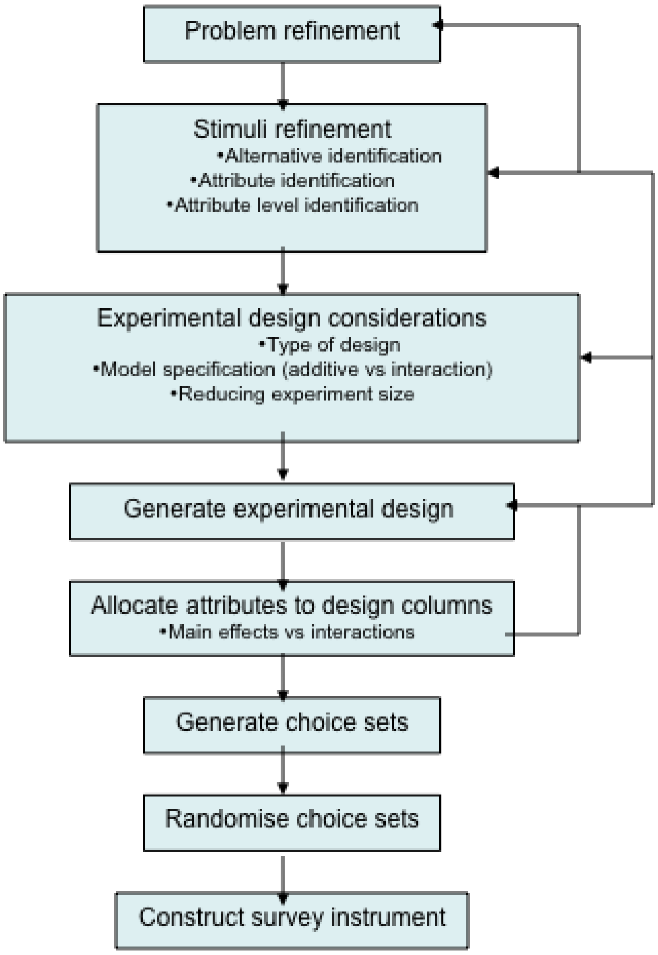


Figure 1: Stages in experimental designs
Source: Hensher *et al.* 2005, p.102.

The first two stages explore the context of the decision and determine relevant alternatives, attributes and levels on which the models will be based. At this stage the modeller is exploring issues that may affect the decision process. When the context is not well understood, a preliminary qualitative study, in the form of a focus group discussion or an administered questionnaire, is required. The aim of the preliminary inquiry is to identify the choice alternatives considered to be relevant to the decision makers and the factors they consider as being influential on their choice.

SP experiments *extend* reality and it is important to work out how far the design can push the limits. The preliminary inquiry should determine what is reasonable in the minds of the decision makers. Whilst it recommended that the range of attributes be wider as it helps achieve efficiency (Rose and Bliemer 2009) this range should not extend beyond credibility. In the market place, the level of service or quality of the alternative is linked to its price. SP experiments need to suspend this relationship, as the aim is to identify the part worth of each quality feature as well as the decision maker's responses to pricing. This is not possible when quality and price are perfectly correlated. The preliminary inquiry seeks to understand the degree in which the price quality relationship can be suspended before the alternative appears 'too good to be true'.

Design considerations and generating the experimental design balance a richer insight into the decision process – achieved by delivering more scenarios to more respondents – and limiting the scope of the survey – managing the cost of the survey and being mindful of the task complexity and respondent fatigue. Experimental designs are similar to sample designs in that the analyst is looking to find the smallest possible sample that achieves a specified accuracy of an unknown population mean (in this case choice parameters). The experimental design population is the universal set of possible tasks – a full factorial design.

At the stage of generating designs the analyst needs to consider:

- the model to be estimated;
- the number of parameters;
- whether the model will include interactions between attributes;
- whether the alternatives should be labelled (car, bus or train) or should be presented as unlabelled collections of attributes (alternative 1 and alternative 2).

The size of the experiment (number of scenarios shown to respondents) is problematic because the universal set can be quite large. Recent designs completed by PATREC include universal sets of size 172 million (Olaru *et al*. 2011) and $5X10^{17}$ (Jabeen et al. 2012). Even when fractional (part of the full set of treatments) sets are chosen to maintain the principle of orthogonality (independent attribute levels), the resulting design is far too big to administer to respondents. The solution has been to optimise efficient designs (Sections 3.2 and 3.3 will expand on this issue).

The size of the experiment – further considerations: Regardless of the approach, the size of the experiments increases considerably with the number of attributes, levels, and alternatives (especially for labelled experiments). Thus it is wise to pay attention to the most relevant elements of the design. Caussade *et al.* (2005) found a substantial impact of the design features on the quality of the data, with the most critical being the number of attributes and number of alternatives:

- the number of attributes/factors had a detrimental effect on the ability to choose, contributing to a higher error variance (the number of levels had a much smaller negative effect); experiments designed and tested by the PATREC research team indicate a maximum of 9-10 attributes per alternative;
- the number of alternatives gave an inverted U-shaped pattern with the optimum at four alternatives; the number of travel alternatives relevant for STEM and ROM24 models currently exceeds four, thus either a sampling procedure should be applied or the conditioning and pivoting should indicate the top four feasible travel modes for the subject;
- the number of choice situations/scenarios/games/stimuli also gave a U-shaped pattern (with an optimum around 9–10); tests for 8, 10, 12 scenarios need to be undertaken in a pilot survey prior to the deployment of the SP survey. The experience with the two ARC Linkage Projects, LP100100436 and LP110201150, suggest that respondents cope well with answering 12 scenarios, if they are explained the task.

Using 66 separate surveys from online panels, Louviere *et al.* (2013) analysed completion rates for discrete choice experiments depending on the number of scenarios, alternatives, nature of attributes, and type of design. They found that choice sets did not reduce considerably the response rates, but the number of alternatives did.

Allocating attributes to design columns is not much more than replacing any effects coding that may have been used with the actual attribute levels and presenting the order in which the attributes will be presented to the respondents[6]. **Generating choice sets** is combining the alternatives (collections of attributes, labelled or unlabelled) in to choice tasks or scenarios. **Randomising the order** in which the choice sets appear is to overcome any primacy effect (tends to process the first in the list) or regency (tends to remember only the last item in the list) effect or to detect any unengaged responses (e.g., always selects first option). Randomising choice sets is much easier using a computer mediated instrument, but is not impossible to do when administering a paper and pencil survey. The presentation of the instrument, as well as an introduction to explain the task and its purpose is not trivial. The instrument should be trialled and piloted before the major release. PATREC experience is about three – five internal trails and 2-3 small pilot tests.

Construct survey instrument refers to creating the pencil and paper form or the online pages for deployment of the SP survey. The mode choice scenarios are typically asking the participants to choose one option (their preferred alternative) of travel. More recently, the Best-Worst (B-W) format has been applied in the SP context. Developed by Louviere and Woodworth (1990) and described by Finn and Louviere (1992), B-W scaling allows for richer information, by eliciting from the respondents not only their stated choice, but also their least preferred alternative and often their second best and worst options, etc. Hence, more information is collected. For a set of three alternatives, B-W provides a complete ranking, whereas with four or more alternatives, a partial ranking can be achieved. This methodology is deemed superior, because it provides more data in an easier fashion (compared to ranking) (as evidenced by studies in marketing and health economics - Auger et al. 2007; Flynn et al. 2007; Cohen 2009; Rose 2013). This approach has shown to lead to more significant parameter estimates and is recommended for the SP survey associated with PARTS 2016.

Once collected, the B-W data could be "exploded" for analysis in two ways:

- By comparing the best option (i.e. the choice) with all the other alternatives, then after removing the best alternative, comparing the remaining options with the worst alternative; this is termed *exploded logit data setup*;
- By creating two choice situations, one with the choice being the best option and another with the choice being the worst option. This is termed a *B-W data setup*.

Prior research found substantial differences between the results obtained with the two data settings, and the latter (B-W data setup) provided additional insights on the decision-making mechanisms surrounding avoidance of unpleasant, risky, uncomfortable situations. Although beyond the scope of this research, PATREC recommends that both data settings should be investigated for PLATINUM development.

3.2 EFFICIENCY

Experimental designs for SP surveys evolved to match the complexity of recent advances in discrete choice modelling. The current state-of-the-art design principles have moved away from the standards of orthogonality, balance and without dominance, in favour of efficiency (Bliemer *et al.* 2009a, b). This is because of the realisation that the traditional principles, in particular orthogonality, were mainly developed for general linear models and not for discrete choice models (Olaru *et al.* 2011).

[6] The experimental designs in this project are optimised considering the absolute values of the attribute levels and not their orthogonal codes.

Efficient designs aim to produce stable and reliable parameter estimates in a fractional design setting, i.e. using only a sample of the possible combinations of attribute levels (Ryan *et al*. 2007). This is achieved by minimising at least one property of the asymptotic variance–covariance (AVC) matrix: determinant, trace, or variances.

Two of the researchers in this project team have designed optimal designs by applying the D-optimality criterion, which is to minimise the determinant of the AVC matrix. This matrix depends on the type of choice model used for estimation, thus the model has to be considered at the time when the SP survey is designed. Rose and Bliemer (2005a,b) and Bliemer *et al.* (2009a) confirmed that efficient designs have more reliable parameter estimates than orthogonal designs for some variants of the discrete choice model.

McFadden (1974) derived the AVC matrix for the multinomial logit model by solving the second derivatives of its log-likelihood function with respect to the parameters. The same method was used to arrive at the AVC matrix for other choice model error structures (see for example, Olaru *et al*. 2011, who presented some of the chronological developments).

One of the challenges present at the design stage of SP is the requirement for a known or previously estimated vector of parameters (priors) when deriving the AVC. A less efficient alternative is to base the design on a set of null parameters (i.e. all equal to zero) or to base the design on the researchers' intuition. However, obtaining reliable prior estimates helps considerably in the design process, because a misspecification leads to losses in efficiency (Kessels *et al*. 2008; Bliemer *et al*. 2009b). Priors may be obtained from literature reviews or pilot studies. Even if the literature matches the research question, it will almost always refer to a different setting. Pilot studies have the advantage of reviewing the survey instrument as well as providing the prior estimates. However, to do these studies well, imposes an added cost in both terms of time and money.

The second challenge with optimised SP designs is that the number of potential combinations increases exponentially with its complexity. In a simple multinomial (MNL) choice model with four alternatives, having three attributes at three levels, the number of potential choice sets exceeds five hundred thousand ((3^3)^4). A residential choice experiment presented by Olaru *et al*. (2011) had over 172 million potential profiles. At this magnitude, a complete enumeration of the combinations of attribute levels is impractical. Some researchers have employed searching heuristics (swapping procedure or the cyclical method) to find efficient choice sets. The D-errors for a large number of possible designs are measured and the more efficient designs are retained. While these represent an improvement to an enumeration strategy, the methods are not classed as optimisation algorithms.

3.2.1 Optimal Designs

Optimality of SP designs goes back to the work of Huber and Zwerina (1996), who described a set of features they believed characterised optimality of choice sets: *level balance* (attribute levels are presented in equal frequency in each choice set); *orthogonality* (attribute levels are uncorrelated to one another and their combinations do not exhibit a certain pattern); *minimum overlap* (for each choice the repetition of an attribute level is minimised); *utility balance* (options within a choice set should have nearly equal attractiveness to subjects). However, as shown by Street and Burgess (2007) these design principles may be met by examples of sub-optimal designs, some of which do not include the main effects. In their earlier work, Street *et al.* (2005) provided strategies for obtaining "quick and easy choice sets" for paired comparison designs when the attributes have the number of levels being equal to powers of two.

Sándor and Wedel (2002) dropped the strict level balance and orthogonality as design principles, and kept minimal overlap in the starting designs in their algorithms. As the authors stated, sacrificing these criteria, created the opportunity for generating statistically more efficient designs, with smaller asymptotic standard errors. The reason behind this is that whereas orthogonality alleviates multicollinearity and minimises the determinant of the variance–covariance matrix for generalised linear models, it plays a minimal role in discrete choice models differ, where only the differences in the utilities of the alternatives matter.

An alternative design principle was then to minimise at least one property of the asymptotic variance–covariance (AVC) matrix: determinant, trace, or variances. Drawing on the development by McFadden (1974), who calculated the second derivatives of the MNL log-likelihood function with respect to the parameters, a substantial number of researchers directed their efforts to minimising the AVC matrix or maximising the Fisher's information matrix (second derivative of the log-likelihood function. This indicator $\{det[I^{-1}(\beta)]\}^{1/k}$), where k is the number of parameters, is known as D-efficiency[7] and has become the standard as they achieve joint statistical efficiency.

Numerous optimisations were suggested, such as: ad-hoc manipulations, heuristics (swapping procedures, cyclical, column/row algorithms etc.), or as in this project, genetic algorithms (Olaru *et al*. 2011). The basic idea is to find experimental points (or profiles) that are ensuring a good coverage of the problem domain, whilst minimising D-efficiency. For simplicity, some designs explored experimental conditions uniformly scattered (Wang and Li 2002).

Street and Burgess (2007) have followed a distinct strategy. An unlabelled sequential method was adopted to generate designs that maximise attribute level differences in each choice set. Street and Burgess (2007) built profiles of the first alternative in each choice set with an orthogonal main effect plan (OMEP), and then sequentially generated profiles for the rest of the alternatives based on some systematic changes in levels derived from the first alternative. Whilst such a design method is useful in locating optimal or nearly optimal designs, it is applied only for simple choice problems and for MNL models.

Currently, the most widely used approach is to maximise statistical efficiency by minimising D-error (Rose and Bliemer 2009). But this approach requires that both the model (Ryan *et al*. 2007) and the parameter estimates (e.g., Ferrini and Scarpa 2007) are known before the SP survey is designed. In applied work, MNL or nested logit are the most commonly assumed choice models (Kaninen 2002; Bliemer *et al*, 2009a; Goos *et al*. 2010) with very few exceptions (e.g., Sándor and Wedel 2005; Bliemer and Rose 2010).

When no information about the parameters is assumed, β=0, then the design criterion is D_0-optimality. Alternatively, a set of non-zero parameters may be introduced into the model based on a pilot study or a literature review. The accuracy of the prior parameters affects the performance of the experimental design and the researchers may have to ponder as to what values these will be. If the aim is to minimise the required sample size, the criterion may consider only the variances from the AVC and not the determinant (Bliemer and Rose 2005; Rose and Bliemer 2009). This is because in a design, the parameter with the highest theoretical required sample size to be statistically significant dictates the sample size for the study.

Numerous algorithms to obtain and implement optimal designs have been examined. Ferrini and Scarpa (2007) and Bliemer *et al.* (2009a) described the construction of various types of experiments including the systematic row- and column-based algorithms (modified Fedorov) and the RSC (relabelling, swapping, and cycling) algorithms. These algorithms start either with a large number of choice sets generated from full or fractional factorial designs, then combinations with lower D-errors

[7] Perhaps, better named D-inefficiency as it calculates the D-error.

are retained (row-based algorithms), or with a randomly generated design in which the levels within each column are changed in order to achieve more efficient designs (column-based); a combination of both strategies such as RSC has also been prevalent. Alternatively, heuristics such as genetic algorithms may be applied. In this case, the chromosomes are the stimuli or choice set elements (combinations of attribute levels across the alternatives), which are combined through various "genetic" operators to obtain the choice sets that minimise D-p (Olaru *et al*. 2011).

Louviere *et al.* (2008) warned researchers that achieving optimality is controversial (heuristics by definition lead to only near-"optimal‟ solutions), plus that different assumptions that researchers make to derive design results are not specifically understood or acknowledged. They advocate for efficiency, as the standard errors of the parameter estimates decrease and hence the parameter estimates are more reliable and attainable with smaller sample sizes (Rose and Bliemer 2005a, b; Bliemer *et al*. 2009a, b). However, the tests performed by Louviere *et al*. (2008) suggested that respondents may become less consistent in answering choice questions as the statistical efficiency increases, which is the "price" that applied researchers have to be prepared to pay.

Although not extensively investigated, another effect of design efficiency is related to the dominant preferences. Results so far showed that optimal designs are less likely to lead to non-trading and lexicographic behaviour (Scott 2002; Hess *et al*. 2010; Olaru *et al*. 2014).

3.2.2 Challenges in achieving efficiency

As indicated, one of the challenges present at the design stage of SP is the requirement for a known or previously estimated vector of parameters (priors) when deriving the AVC. A less efficient alternative is to base the design on a set of null (zero) parameters (i.e. D-z) or to base the design on the researchers' intuition. However, obtaining reliable prior estimates helps considerably in the design process, because misspecifications diminish efficiency.

The second challenge with optimised SP designs is that the number of potential combinations increases exponentially with its complexity. In a simple multinomial (MNL) choice model with four alternatives, having three attributes at three levels, the number of potential choice sets exceeds five hundred thousand ((3^3)^4). Wang *et al.* (2014) obtained 4.78 million fare combinations in an airline revenue management problem with four alternatives and three attributes (price, flexibility, and time). The current study in to Perth's Park and Ride (PnR) function includes a mode choice and time of say SP experiment with $3.9X10^{11}$ possible alternatives in the universal set[8]. The D-errors for a large number of possible designs are measured and the more efficient designs are retained.

Olaru *et al*. (2011) have designed optimal SP designs using genetic algorithms (GA). Within the Transport Research Group at UWA, GAs have been successfully applied for a considerable number of various combinatorial problems in transport (see Taplin *et al*. 2005) and the research team members have confidence in the power of GA to offer efficient design solutions.

The third issue is deriving the AVC. If the interest is to obtain more flexible structures, capturing the preference heterogeneity, mixed logit (ML) models could be applied. In ML, both the parameters in the utility functions and the error terms are random variables. ML are superior to MNL and NL, however pose an additional challenge in deriving the Fisher Information Matrix. Bliemer *et al.* (2010) have shown that the MNL designs are robust to ML modelling.

---

[8] Unpublished work undertaken by Huang Ying for her PhD studies at UWA Business School.

3.3 SAMPLE SIZE COMPUTATION

The efficient design principles permit a small fraction of the full factorial set of attribution combinations to be used. The idea is to maximize the design efficiency by finding the mixture of scenarios that produce the smallest standard error for the set of parameter estimates. Two calculations are relevant:

1) D_p: the determinant of the Asymptotic Variance-Covariance Matrix, which is derived from the Fisher Information matrix I (its inverse over all respondents); the p stands for with **prior** estimates. The metric gives an overall measure of efficiency of the design matrix; taking into account the individual parameter standard errors (along the diagonal of the matrix) and the correlations between the parameters (the off diagonal elements of the matrix):
   $\{det[I^{-1}(\beta)]\}^{1/k})$ (1)
   The goal is to find the minimum determinant, D_{pmin}, for a pre-determined number of scenarios to be completed by each respondent. The sample size is computed by finding the smallest sample that ensures that all parameter estimates are significant at the required level of significance.

2) S_p: the corresponding sample size to the parameter with the largest standard error. The objective is to find the smallest S_p for which all parameters are significant at the required level. This is equivalent to finding the sample S_{pmin} that ensures statistical significance for the *least* reliable parameter with the smallest sample size.

The experimental designs presented in Section 4 and in the EXPERIMENTAL DESIGN TECHNICAL APPENDIX are optimised for MNL models using GA (using Palisade Evolver software).
Both optimisation of the D_p and of the minimum sample size (S) required is applied. This is important, given the increased costs of surveys.

The sample size calculation is based on the ratio between the prior parameter value and its standard error generated by the choice of design. The number of surveys to be completed, M_{max}, is the smallest sample size required to provide significant parameters in a choice model using the prior information. For each parameter the required sample M_k is found by setting the asymptotic t ratios to the desired level of significance (1.96 is appropriate for the 5% level of significance) and solving the following formula:

$$\frac{\beta_k}{\frac{se(\beta_k)}{\sqrt{M_k}}} = t_k \qquad (2)$$

The sample size is calculated for each parameter in the experimental design using the diagonal elements in the asymptotic variance-covariance matrix. The largest $M_k = M_{max}$ corresponds to the '*least certain'* parameter (the parameter with the highest coefficient of variation). Consistent with prior scholarly work, results in Appendix C show that experiments that are S optimised require smaller sample sizes than D optimised designs.
Even though the literature is ambiguous about the choice of objective function for optimised designs, it appears that D-error is crucial for designing experiments tor testing behavioural models where correlations between attribute levels and interactions are present. Where the off-diagonal terms are of little concern and the analyst seeks only reliable parameter estimates for ex post evaluation calculations, the S optimisation is more appropriate.

3.3.1 Representativeness

Literature in data collection (regardless of the area of interest) is clear about the types or errors affecting representativeness and their magnitude (Richardson *et al.* 1995). Simply said, a sample is representative of the entire population if: it is “similar” to the population from which is drawn and it covers the population heterogeneity (space, socio-demographics, behavior of interest). This draws attention to the fact that a sample representative for a particular question/matter is most likely not representative for a different question.

Stochastic sampling procedures ensure representativeness, however non-response and response biases can have a large impact on the validity of the questionnaire or survey to which the participant is responding. This is because human subjects do not respond passively to stimuli, but rather actively integrate multiple sources of information to generate their response in a given situation. If they are disengaged with the survey, they may refuse participation (non-response). On the other hand, high involvement with socially and morally charged issues may cause not only self-selection, but also bias in responses (denying undesirable traits, and ascribing to those that are socially desirable). Both sources create a systematic error of the measurement that is sought (Bonsall 2009).

More recently, representativeness was questioned in relation to the method of data collection. Some deployment techniques are considered more immune to such biases: pencil and paper postal surveys, face-to-face interviews. However, empirical knowledge concerning the validity of data gathered using new digital communication technology approaches (GPS, web surveys), in comparison to more traditional surveys, is still in its relative infancy. This suggests that additional attention must be paid to any link between the survey instrument (including SP), the medium of data collection, and the non-response and social desirability biases. For example, relying solely on socio-demographics for representativeness may obviate obtaining reliable findings if ignoring destination locations for activities or mobility restrictions.

Achieving representativeness and minimising biases is crucial for transport planning policy considerations, as formulation of effective interventions requires valid data from all households or individuals, accounting for the types who often do not make it into our samples.

3.4 BAYESIAN DESIGNS

Recognising the importance of the priors for the quality of the designs, more recently Bayesian designs have emerged on the arena of experimental designs (Kessels *et al*., 2008). Bayesian approaches go beyond the appealing assumption of fixed known parameters and take into account their uncertainty. This requires efficiency to be calculated as the expected value over numerous draws taken from the probability distributions of the priors (Bliemer *et al*., 2009b: 100) and the resulting design is called D_B-optimal. For a better coverage of the distribution space of the distributions for prior parameters, quasi-random Monte Carlo simulation or Gaussian quadrature are recommended. Bliemer *et al*. (2009b) found better approximations to the true level of efficiency of a design using the Gaussian quadrature methods.

If requested by the Working Group, examples of Bayesian designs can be further explored in 2015.

4 STATED PREFERENCE FOR PERTH METROPOLITAN AREA

This section presents a number of policy questions for which the SP survey can be designed to address pressing questions identified for the choice models currently included in STEM and ROM24.

4.1 SP METHODOLOGY RELEVANT TO THE HOUSEHOLD TRAVEL SURVEY

The primary purpose of the SP survey is to provide parameter estimate for Perth's strategic transport models. This is in response to a gap in the current mode choice components of the existing strategic models:

> *"A major need reported in the Needs and Gaps report was to make the mode choice model more responsive to fares and wait times*" (Taplin *et al.* 2014, p.46).

It was also recommended that SP is the appropriate tool for filling this gap:

> *"The new choice parameter estimates should be made by the advanced stated choice (SP) methods that have been developed in the Australian Research Council Linkage project LP110201150, in consultation with Professors Hensher at Sydney University and Rose at the Institute for Choice, UNISA".* (Taplin *et al.* 2014, p.48).

The review went on to indicate that the mode choice models explore the factors beyond travel time and travel cost:

> *"To achieve a fully responsive choice model it will also be necessary to model additional factors. Travellers are increasingly sensitive to crowding on public transport, to highly variable road travel times and to parking availability. Thus a satisfactory mode choice module should contain parameters reflecting responses to these factors, as well as the time and cost parameter estimates already included".* (Taplin *et al.* 2014, p.48).

The ARC supported "Modelling and evaluating the joint access mode and train station choice" project (ARC LP110201150) provides preliminary information that travel time variability and uncertainty about parking availability play a significant in commute mode choice. These effects are detected in both the RP and SP data. However, the data suggest that crowding does not influence mode choice.

Recommendation 1 A stated preference instrument is to form part of the household travel survey. A base design is developed with the sole purpose of providing parameter estimates for the mode choice models in Perth's strategic model (currently STEM and soon to be PLATINUM). Furthermore, the SP instrument should include attributes that are not currently expressed in the STEM mode choice models. Specifically, response to travel time variability and parking availability should be investigated.

Commuting (journey to work or education) places the most stress on the transport infrastructure of Perth. Significant policy decisions will require the support of a model of travel behaviour for this function more so than other travel purposes. The review identified gaps in the mode choice model for shopping trips (Taplin *et al*., 2014; p.47). Although the SP can be applied for any trip purpose, the experimental design templates offered in this report focus on commuting/education trips. This is because the DoP Steering Committee agreed that the survey investigate specific policy areas (See Section 4.2 below) and these relate to a.m. and p.m. peak travel.

Recommendation 2 The Stated Preference component to be applied to the journey to work or education travel purpose.

4.2 POLICY FOCUS

The SP surveys designed in this project look into mode choice. Each experiment will include commonly specified attributes in mode choice models: public transport headway, fares, price of fuel, waiting and transfer time. However, to add value to the metropolitan survey, the SP component will examine a policy as determined by state government authorities. More than one issue can be explored, but it is likely to require a more complex task or a portioning of the sample so that one sample responds to one task and another sample tackles a separate issue. During the Steering Committee meeting on Oct-17 2014, it was agreed that a **base design** and **two policy areas** be explored further:

- **Managing congestion during peak hour -** Managing travel demand on limited road supply during peak hour including the use of financial instruments;
- **The demand for a "new" public transport mode** – Forecasting the patronage for a new light rail service.

Exploring the two policy areas of congestion during peak hour and a new mode is meant for demonstration and does not lock in any particular policy investigation in the main survey. Other policy issues raised in the Steering Committee meeting were:

- **Parking -** benefits of PnR, parking charges and regulation;
- **Public transport demand**- managing travel demand on a limited public transport supply during peak hour;
- **"New" public transport operations** – flexible transport solutions (demand responsive transport);
- **Mono modes –** exploring factors that affect the choice of a minor mode in more detail (e.g., cycling).

Recommendation 3 PATREC recommends that two or three key policy questions are identified and that the SP instrument is designed to provide demand insight to support decisions in these policy areas. However, the experiments should be extensions to a base design aimed at achieving ***Recommendation 1.***

The PATREC advisory team believes that extending the base SP design to investigate demand for a specific research question poses little risk in undermining the primary purpose of the SP component (***Recommendation 1***). However, pilot studies should be undertaken using all experimental designs to test the parameter consistency across the choice models. If it is found that one or more of the policy specific designs diverge from the base design, either adjustments to the designs is required or a greater sample should be allocated to the base design.

Recommendation 4 Supporting the mode choice models in Perth's strategic model has a higher priority than exploring specific policies. A substantial pilot test should be conducted on the survey instruments to judge the quality and consistency of the designs, due to their strong dependence on prior parameter estimates.

4.3 COMBINING RP AND SP

Taking into account the priority of ***Recommendation 1***, the SP component of the survey needs to be solidly embedded into the main survey instrument. Section 2.2 of this report provided a review of the state-of-the art methods for combining RP and SP data into a single model of travellers' choices. Data enrichment primarily refers to the modelling phase, however it informs the experimental design (Section 2.3). It is essential that the survey adopts best practices at the design stage, during data collection and when developing the models.

Data enrichment: describes the joint estimation of behavioural parameters using both RP and SP observations. The methods of Basic and Advanced Enrichment are detailed in Section 2.2. A third method is discussed, but this ignores the parameter estimates from the RP data (other than to arrive at the scale difference between RP and SP). The modelling strategy should pursue data enrichment and this informs the experimental design and setting of attribute levels.

> ***Recommendation 5*** Future mode choice components within Perth's strategic model should aim to make the most of data enrichment to provide robust parameters.

Data enrichment informs the general principle of grounding SP attributes by the current experiences of the respondents (i.e., the revealed preference data). However there are a number of approaches to ground the SP attributes.

1) **Pivoting, using a reference alternative**: The current RP choice with its observed attributes appears in each SP scenario. The SP alternatives share these attributes but have levels that are proportionally higher or lower (e.g., by 25%). The choice task becomes one of 'would you switch from your current alternative?'
2) **Pivoting, using an Experienced Alternative**: This is a less strict form of referencing is to pivot all SP attribute levels around an experienced alternative. The difference from the referenced alternative approach is that the attribute levels of the experienced alternative are included in the design.
3) **Conditioning the SP attributes levels**: Conditioning the SP alternatives further relaxes the constraints of presenting the current alternative in each scenario. Where no RP data is collected it is recommended that network data be reviewed to ensure that the SP attribute levels reflect the current travel conditions. Where RP data is collected from a different sample, SP attributes are grounded by using the average levels for the relevant segments of the market.
4) **Conditioning the SP choice sets and attributes levels – a hybrid approach to pivoting and conditioning:** The respondent's current alternative (RP) appears in each choice set, but with attributes conditioned by the average levels for commuters with similar trip lengths (short, medium and long). The strategy requires far less than one design for each respondent, but more realistic than ignoring the respondent's current choice and context.

Pivoting on a reference alternative is unsuitable for this SP survey because the sensitivity to attribute levels for the current mode is important. Pivoting on an experienced mode is relevant but represents one extreme of needing to design scenarios at the level of the individual. Conditioning on attribute levels only is an inferior choice and only applies to situations when RP data are not collected. The Hybrid approach is adopted in the ARC project LP110201150 (Modelling and Evaluation of the Joint Access Mode and Train Station Choice, PnR) and has yielded sound estimates of parameters.

Recommendation 6 The SP choice sets and attribute levels should be conditioned on the RP data using an appropriate segmentation that takes into account the respondent's context. This may be the length of the trip or it may take into account other aspects, such as the place of work being within Perth CBD.

4.4 ADDITIONAL CONSIDERATIONS

4.4.1 Attitudinal Scales

Preferences and psychological strategies aimed at changing attitudes have been found influential and hence, have recently gained prominence in the transport policy agenda. Programs such as TravelSmart (Department of Transport, Western Australia 2010, 2014) or similar individualised marketing and travel feedback programs are regarded as effective, low-cost solutions for urban transport. This means that changes in attitudes can be achieved with persuasion, but first, preferences and attitudes need to be known and understood.

Substantial scholarly work has shown that inclusion of attitudes in discrete choice models (hybrid choice models) explain better individual decisions of mode (Ben-Akiva *et al*. 2002; Schwanen and Mokhtarian 2005; Vredin Johansson *et al.* 2006; Ettema *et al.* 2011; Prato *et al.* 2012; Atasoy *et al.* 2013; Van *et al.* 2014) and relocation choice (Walker and Li 2007; Walker *et al.* 2011; Olaru *et al.* 2011), among others. This is because a house or a car is not appreciated only because of its instrumental functions. Besides, other motives seem to play an important role: feelings of power, comfort, social status, restrictions, etc., which implies that the utility functions should embed not only features representing the instrumental value of the alternative, but also attitudes and feelings of the decision maker.

Many studies have used different methodologies to examine attitudes toward car, public transport, and active travel. In essence, these attitudes refer to three main components: symbolic (how people express social and personal identity), instrumental (functional attributes and benefits of using transport modes), and affective (emotional feelings of travellers) (Van *et al.* 2014).
They can be incorporated either as covariates indicating inclinations/predispositions of the respondent, or perceptions/attitudes towards characteristics of the alternative, influencing the individual choice. Appendix B exemplifies statements for both approaches.

Acknowledging the role of attitudes in disentangling travel decision processes, PATREC recommends:

Recommendation 7 Both the RP and SP surveys make use of attitudinal scales to improve the estimation of choice parameters and to inform behaviour modification policy.

4.4.2 The Transition from STEM to PLATINUM

To support the current strategic models, the SP needs to be trip-based. However, a responsive behavioural model implies a replacement of the trip-based modelling approach by a tour-based schema. Given the complexity of a tour-based SP, expected to increase considerably the number of alternatives[9] and the attributes required, at this stage we recommend a mid-way solution, with the SP survey only including possible constraints/activities in the afternoon that may affect the modal choice for their morning commute *("Which of the following activities do you undertake on your*

[9] For each commute alternative in the morning, there are potentially many travel alternatives for returning home. Moreover, the tours should consider the number of activities included and their timing/sequencing.

return trip?") and relevant attitudinal questions (examples of indicators of flexibility and convenience are provided in Appendix B).

4.4.3 Time of Departure

Successful implementation of the time-of-day modelling capability would assist STEM and PLATINUM to produce and refine the 'n-hour' O-D matrices to cover the hours of the day, potentially incorporating factors for peak spreading. This report suggests two solutions to address the issue: a) including the departure time as an indicator for alternatives; and b) using time-of-day as a blocking factor. The former may create potential issues with the realism of the scenarios (e.g., departing off-peak and facing worse traffic conditions than during peak time), whereas the latter disconnects to a certain degree the departure time from the choice of mode in the scenario (all alternatives in a scenario assume about the same time window to reach the destination). Nevertheless, blocking allows for exploring substantial changes in departure time, depending on the individual's circumstances, which could inform policies for peak spreading. For the SP survey included in PARTS 2016, PATREC recommends applying time-of-day as a blocking factor.

Recommendation 8 The SP survey should account for departure time and incorporate it in the data collection as a blocking factor.

4.5 EXAMPLE DESIGNS

To address the two policy areas proposed by the Steering Committee (Section 4.2), a baseline design was undertaken first. In line with research that examined the impact of several dimensions of complexity (number of alternatives, attributes, type of design) the designs included in this report are simplified and they do not include all modes and multimodal combinations. For illustrative purposes, we provide designs with four modal alternatives: car (private) transport only, public transport only, active travel, combination of private and public (e.g., park-and-ride). They can be relatively easy to be adapted to reflect the mode segmentation in the transport model and insert combinations of alternatives seen by the respondents as relevant to them (i.e., pivoting the design on the respondent's current revealed choice).
The implications of a more refined choice design for the delivery of the survey and on the required number of responses are discussed after the presentation of the designs.

This section provides a list of potential attributes for each policy question, along with the recommended sample size to meet the statistical efficiency requirement (Appendix C). The EXPERIMENTAL DESIGN TECHNICAL APPENDIX outlines the detail for the designs, the way in which the sample sizes are arrived at and an illustration of the survey instrument for each design is provided in Appendix D.

4.5.1 Base Design for Mode Choice

The SP component of the survey aims to provide robust, but policy sensitive, parameters for the mode choice module of Perth's strategic transport model (currently STEM, but to be PLATINUM in the foreseeable future). To this end, any policy specific SP inquiry should be an extension to a base model that meets the needs of the journey to work/education mode choice model in the strategic travel model. We suggest the Steering Committee to consider a mode choice and time of departure model, as this will quite possible form the basis of the journey to work choice model going forward.

As indicated, the base model outlined in Table 2 includes the four main modes: car, bus/train, bicycle, and combination car-public transport.

Table 2: The scheme for the Mode Choice SP experiment

Focus of the design	Attributes	Travel modes
Base Design: Mode Choice (labelled experiment), with alternative specific parameters	Attributes in this design: • In-vehicle travel time, • Travel time variability, • Waiting time, • Number of transfers, • Travel cost or fare, • Parking cost at the destination, • Access and egress times, • A binary attribute that could refer to departure time before am peak or during am peak, driving alone or with passengers, etc. • Presence of a dedicated bicycle lane, • End of trip facilities. Adjustment to departure time, travel distance, and travel purpose are used as blocking factors. Further considerations include: • Crowding, • Probability of getting a seat, • Wi-fi availability.	• Car • Public transport (Bus/Train as applicable) • Active transport (Bicycle/Walking as applicable) • Combination (PnR, BnR)

The set of parameter priors and resulting designs are presented in the EXPERIMENTAL DSIGN TECHNICAL APPENDIX. The prior parameters are taken from the estimation results from LP110201150 and from a literature review, where the parameters were not part of the prior study (i.e., active transport). The attribute levels relate to morning peak trips of approximately 20km and the values assigned to parameters on waiting time, travel time, and access/egress time suggest a value of time that ranges between \$10 and \$15/hour.

The experimental designs are optimised for MNL models using GA (using Palisade Evolver software). The key findings from the optimisations are presented in Table 3 below. Two sample sizes are calculated. Firstly, the sample size resulting in minimising the determinant of the ACV matrix, D_{pmin}, is roughly 1,200. This is substantially higher than the minimum sample size S_{pmin} = 178. The S_{pmin} is appropriate for main effects models and does not account for possible interactions between attributes in the utility functions. However, the small sample size does represent a substantial cost saving[10].

Table 3: Summary mode choice SP experiment features and sample sizes

Orthogonal design	Dp-optimised design	S-optimised design
D_p = 30.147 S_p = 10^9	D_{pmin} = 8.296 S_p = 1,194 -required sample size The design is relatively balanced. Each attribute level occurs on at least three scenarios out of the set of 12 (in a balanced design each three level attribute will occur exactly four times).	D_p = 11.483 (10.762) Sp_{min} = 178 -required sample size Less balanced. Each attribute level occurs at least twice.

[10] The efficiency criteria ignore representativeness. Despite S_{pmin} indicating a small sample is statistically efficient, this sample would be unlikely to meet representative requirements.

4.5.2 Light Rail Mode Choice

The design for the Light Rail is based on the baseline SP design, and it includes the new option along with the current public transport, with private travel and combination private-public transport (Table 4). The attributes and prior parameters for the alternatives are provided in the EXPERIMENTAL DESIGN TECHNICAL APPENDIX.

Table 4: The scheme for the Light Rail SP experiment

Focus of the design	Attributes	Travel modes
New public transport alternative: light rail = Base Design + attributes of new mode	All attributes from the base design • In-vehicle travel time, • Travel time variability, • Waiting time, • Number of transfers, • Travel cost or fare, • Parking cost at the destination, • Access and egress times, • Presence of a dedicated bicycle lane, • End of trip facilities. Adjustment to departure time, travel distance, and travel purpose are used as blocking factors.	• Car • Public transport (Bus/Train as applicable*) • **Light rail** • Combination (PnR, BnR)

Note: * The current public transport option will be set depending on the location of the suggested new LR line (e.g., for LR to the Northern suburbs, the train to Butler could be an alternative, for "Knowledge link" LR, the bus is deemed as the alternative).

The required sample sizes given in Table 5 are not so far apart as was the case in the base design. The sample size resulting in minimising the determinant of the ACV matrix, D_{pmin}, is roughly 750 and the minimum sample size, S_{pmin}, is approximately 550. The almost double sample size is a result not only of the different alternatives and relevant attributes, but also of the way restrictions are embedded in the optimisation. The base design regards only moderate balance and there is no indication of near-orthogonality, whereas the light rail SP design also requires correlations between attribute levels to be limited to an absolute value of 0.4.

Table 5: Summary light rail SP experiment features and sample sizes

Orthogonal design	Dp-optimised design	S-optimised design
D_p = 72.94 S = 10^10	D_{pmin} = 15.725 S_{min} = 743 Less balanced compared to the base design	D_p = 19.813 S_{min} = 541

4.5.3 Car Toll Mode Choice

The design for the Car Toll is based on the baseline SP design, and it includes the new option along with the current public transport, with private travel and combination private-public transport (Table 6). The attributes and prior parameters for the alternatives are provided in the EXPERIMENTAL DESIGN TECHNICAL APPENDIX.

Table 6: The scheme for the Car Toll SP experiment

Focus of the design	Attributes	Travel modes
Managing Congestion = Base Design +	All attributes from the base design • In-vehicle travel time, • Travel time variability,	• Car • Car toll • Public transport (Bus/Train as

Instrument and uncertainty	• Waiting time, • Number of transfers, • Travel cost or fare, • Parking cost at the destination, • Access and egress times, • Presence of a dedicated bicycle lane, • End of trip facilities. and the instrument: • Toll (undefined) Further considerations may include: • Congestion Charge • Infrastructure toll or route specific toll • Managed Freeway on ramps • Time spent in congestion Adjustment to departure time, travel distance, and travel purpose are used as blocking factors.	applicable) • Combination (PnR, BnR)

The design for congestion management through the use of a toll requires several additional constraints:

- The public transport fare and the egress time are considered the same for the Public transport only and for Combination car-public transport;
- Parking cost at destination is the same for the two private transport alternatives car and car on a tolled route;
- Car travel time and associated fuel cost must be larger than the travel time and fuel cost on the tolled route; similarly for variation of travel time.

These constraints affect the efficiency of the design and consequently, require larger sample sizes. The required sample sizes given in Table 5 are not so far apart as was the case in the base design. The sample size resulting in minimizing the determinant of the ACV matrix, D_{pmin}, is roughly 750 and the minimum sample size, S_{pmin}, is approximately 550.

Table 7: Summary Car toll SP experiment features and sample sizes

Orthogonal design	Dp-optimised design	S-optimised design
D_p = 31.72 S = 10^9	D_p = 6.794 S_{min} = 1,564	D_p = 9.423 S_{min} = 959

PATREC recommends the use of templates for the SP survey associated with PARTS 2016, either in their current form, or amended by changing appropriately attributes.

> ***Recommendation 9*** SP surveys can use the three templates provided by PATREC or, if necessary, customise them by changing attributes with similar scales/ranges and parameter estimates.

5 SURVEY ADMINISTRATION AND ASSOCIATED COSTS

5.1 SAMPLING AND RECRUITMENT

A simple random sampling is usually straightforward to implement, but other kinds of sampling can give better estimates. Stratified or cluster random sampling partitions the population into groups called strata/clusters, and then randomly samples within each group. However, a sampling strategy that is optimal for estimating one type of parameter may be inefficient or even unusable for another, so there is often a compromise between various procedures (stratification, pure random sampling), so that useful information can be obtained about a wider set of parameters.

A multi-stage sampling – starting with suburbs/areas within the Perth metropolitan area, then probabilistic sampling within areas is envisaged for the SP-enhanced PARTS survey. This approach will ensure spatial coverage and consistency with the socio-demographic profile of the population.

Data Analysis Australia indicated three possible strategies for the main (RP) survey. For all three the initial recruitment of the household is a face-to-face contact. However, the degree of contact varies after this initial contact. The face-to-face option (highest cost) includes a household visit to collect the travel diary. It appears that the validation of the GPS data is to include a face-to-face interview. However, the method of 'prompted recall' is not spelt out in their report. If the prompted recall includes a face-to-face interview, then this would greatly assist the collection of the SP data. Finally, a pencil and paper survey will put time pressure on the SP team to recontact the household, because the survey data would need to be entered into a database before appropriate analysis can be undertaken.

The sampling strategy, method of recruitment and the primary survey instrument will be in the remit of the main survey agency and, as such, the issue is not further explored in this report. The instruments below list the possible ways of administering the SP survey and outlines their relative benefits and shortcomings.

The sample sizes given in Section 4 represent the basis of a cost estimate given in Section 5.4.

5.2 SURVEY INSTRUMENTS

5.2.1 Face-to-Face Computer Aided Personal Interview: Secondary Provider

In-depth personal interviews may be more appropriate given the complexity level of certain SP surveys, but these are the most costly methods of data collection. They have the lowest refusal rates and are likely to lead to the lowest level of response errors. Face to face interviews enlist visual aids and commentary to explain the task. A computer assisted personal interview (CAPI) has the advantage of being dynamic in that the choice context can be delivered based on the respondent's revealed choice. Additionally, a CAPI automates data entry and checking and has the potential to geocode automatically if a map-based interface is used.

Sampling frame being PARTS 2016 respondents

This is the highest cost option explored by Data Analysis Australia. However, it does provide a distinct advantage to the collection of quality SP data. The CAPI tool developed for the RP collection could provide a seamless transition into the SP component for a sub-sample of households. However, the calculation of real time choice sets and attribute levels represents a substantial programming component. It would be too risky to compute efficient designs "on-the-fly" and some intermediate conditioning/pivoting process using a databank of designs would need to be used.

5.2.2 Integrating with the Prompted Recall

Forming Part of the GPS Prompted Recall, Face to Face

In-depth personal interviews (SP) are conducted at the same time with the RP, as the household is being asked to validate (GPS) or recall (travel diary) their trips. The face-to-face interview will have a low refusal rate, because the respondents have already agreed to take part in the survey. Integrating the SP data collection into the RP validation/recall interview represents a saving on the marginal cost of the SP component. One advantage for this approach is that there is some lead-time to condition the SP survey on the RP choices made by the household.

Forming Part of the GPS Prompted Recall, Telephone

The survey would lose the personal touch and the capacity for the interviewer to guide the respondents through the instrument. However, many of the aspects of a face-to-face interview are maintained. A web-based tool could be accessed by the household to complete the GPS validation as well as the SP component of the survey.

Sampling frame being PARTS 2016 respondents

By integrating the SP component with the main PARTS 2016 the respondents will have agreed to be recontacted (this is necessary to meet Australian Privacy Principles). A mail out survey to these households will yield a much higher response rate. The self-administered survey return system used in previous PARTS survey had a 50%-60% response. If a similar method is used, the SP component would expect a return of lower than 50% (due to attrition), but not as low as a first contact by mail only. The paper and pencil option will mean a delay between the primary data collection and the follow up SP survey. A well-organised business process could reduce this turnaround to four weeks. However, delivering the SP component is outside of the main agency's contract, then this turnaround could be substantially longer and could affect the quality of the SP data.

5.2.3 Mail out (pencil and paper or online)

For covering a large geographical area, mail questionnaires (reply-paid) may be relatively economical to implement. However, experience with the "Analysis and modelling of driving patterns for limited range electric vehicles, EV (ARC LP100100436) and the "Modelling and evaluating the joint access mode and train station choice" (ARC LP110201150) suggests a response rate of 10%-20%. The low response rate may also cause issues with representativeness.

5.2.4 Low Cost Options

Online panels, such as PureProfile or SurveyMonkey, ensure data collection of the required sample in a reduced time frame, however the representativeness of the sample to the population is difficult to support. Moreover, the standard panel surveys are not tailored for discrete choice experiments hence require supplementary programming work. Finally, newer methods such as SMS, may offer wider coverage and higher response rate (simpler, quicker, more attractive approaches), but are limiting in the type and size of the questions posed to the respondents (e.g., a standard single SMS text is limited to 160 characters). In Australia, SMS Tech offers competitive rates for web-based SMS surveys (www.smstech.com.au), which are suitable for certain types of travel/activity questions or for collecting panel data.

Online panels are not appropriate for undertaking a survey that requires a representative survey. The panels may be used for a preliminary testing of the instrument. However, the investment in programming would be better spent on developing a computer aided personal interview (CAPI) tool and testing the instrument on a convenience sample.

The decision on whether to use personal interviews, mail questionnaires, or online panels, occurs on the basis of some combination of objectives of the research question, level of accuracy, cost, and time frame available. PATREC recommends that the option to integrate the SP component into the main RP survey. This makes sense in terms of having a consistent channel of communication as well as cost savings.

> ***Recommendation 10***: To aid the consistency of communication with the respondents the SP instrument should be as seamless as is possible, in that it looks and feels part of the main survey. The exact method for administering the SP instrument depends on the choice of instrument made by the survey agency administering PARTS 2016. However, a computer-based tool should be used.

5.3 IMPLICATIONS FOR INTEGRATING WITH PARTS 2016

With the aim of SP being to provide enriched parameter estimates for Perth transport models, it would appear that a full integration with PARTS 2016 is necessary. The RP component will be collected by the main survey and a follow-up survey would be administered to a representative sample. Cost implications may mean that only a fraction of PARTS 2016 sample will receive the follow-up SP survey.

An interim level of integration is possible in that the PARTS 2016 revealed choice attribute levels provide a foundation for future SP exercises. PARTS 2016 will provide a range of attributes that can at least be considered as reasonable – or likely to have been experienced. The parameter estimates from RP choice model will provide priors to be used for efficient designs (aimed at saving on sample size requirements). Conditioning is not the same as pivoting, as the attributes for the reference/experienced alternative are not collected for that individual.

Any SP exercise that is to run independently to PARTS 2016 and aims to support policy will need to incorporate its own RP preliminary survey.

1) **Organisational boundaries**. The stated preference function relies heavily on access to RP data in a time fashion. Cross organisation boundaries will inhibit the sharing of data. The primary surveyor will have a privacy contract with the sample that will require some form of data de-identification to be undertaken before sharing. Along with necessary data cleansing, this will cause delays particularly because the priorities of the RP data collection agency will focused on the efficiency of its primary function and sharing data will require shifting resources away from its core business. By contracting the same organisation to undertake the RP and SP component it is understood that a number of data analysis functions are necessary to complete the full data collection. However, this poses some risk in finding the right contractor with the capabilities to understand both the RP travel survey data collection and the pre-analysis required to develop and administer deliver an efficient and relevant SP instrument.
2) **Cost of administration:** An integrated RP and SP survey undertaken by one agency has the potential to reduce the cost of administrating both survey components. The obvious reason being that only one source of overheads applies to the individual survey components. However, given that the data collection agency will need to follow up with households to verify travel diary or GPS records, the potential saving comes from incorporating the SP data collection into the follow up validation process.
3) **Continuity of conversation with the respondent:** The respondents will view the whole exercise more favourable if there is a consistent channel of communication. The implication for the administering organisation is that they assign a single person to each household to

be the main source of contact. However, from an integration of the SP and RP context, having a single organisation (branding, contact addresses, newsletters, and hearing complaints) is less frustrating to the respondent and will improve the quality of the data.

Recommendation 6 suggests that the contracted survey agency will have the capabilities that span travel survey collection, choice modelling and efficient experiment design. PATREC believes that this is unlikely to be the case and that the Departments of Planning and Transport must consider how it will arrange a contract that ensure these capabilities are brought in-house or to develop a joint venture – a temporary project organisation that includes resources from two or more firms. The other possibility is that the contract with the main survey agency specifies the data preparation and sharing requirements. The advantages and disadvantages of in-house, joint venture or contracted sharing expectations are outside the scope of this report, but these should be carefully considered by DOP/DOT.

Recommendation 11: DOT & DOP seek further information on the best organisational structure to meet the requirements of a joint RP and SP data collection exercise. This may form part of the tender process, whereby potential contractors will suggest ways that they can meet the capability requirements.

5.3.1 Timing between RP and SP collection

A significant advantage of efficient designs is that a high degree of parameter efficiency may be achieved for relatively small sample sizes. In practical terms efficiency improves the quality of the data for less cost. The highest efficiency can be achieved by basing designs on accurate prior estimates as well as conditioned attribute levels and choice sets (see Section 0.3 above). However to achieve the efficiency gain a quick turnaround of the RP data needs to be achieved. This is because a series of choice models will need to be run to identify the appropriate prior estimates for the current sub-sample and conditioned choice sets and attributes will need to be generated based on RP observations. UWA business School has achieved a 12-week turnaround, but this because the process was manual and only person was involved. It is expected that the data cleansing, formatting, choice modelling and design could be hastened by developing a process and automating components of this process.

Recommendation 12: A business process for extracting, cleansing, formatting, modelling and design be created to make sure the SP component is ready to go when validation takes place. This process will depend on the choice of organisational structure as outlined in ***Recommendation 10***.

6 CONCLUSIONS

This report analyses how stated preference (SP) data can become integral to PARTS 2016, in an attempt to enhance the parameter estimates of the current strategic models.

As revealed preference (RP) data is limited to the current technology frontier and market conditions, it cannot inform on contexts that do not currently exist and cannot be responsive to policy changes. Drawing on state-of-the-art practices in designing SP experiments, a number of recommendations are made. They refer to integrating SP in the upcoming PARTS 2016 (and creating the required organisational structures), to conditioning the SP survey on respondents' circumstances (RP data), and applying the SP templates developed for baseline and two policy options for the analysis of commuting travel. The large scale of the undertaking along with the possibility that the RP survey may be conducted using new technology suggests a degree of risk. It is important that DOT/DOP appoint a project management team that oversees the benefit realisation. The client-side project management is critical for a successful data collection.

> ***Recommendation 13***: DOT/DOP investigate the roles and performance tasks of client-side project management in order to arrive at an informed decision on allocating resources to this function.

Experimental designs were optimised with two criteria: determinant, D_p and sample, S. Although the 30-min run of the genetic algorithms cannot guarantee the optimum was achieved, we are confident that we found "good designs", able to maximise reliability, minimise correlations of parameters, and ensure relative balance. The S designs are much more efficient than the D_p designs, and both outperform the orthogonal designs.

It is recommended that the contracted survey agency has these skills in-house or is willing to acquire these resources. The RP and SP surveys should be administered or, at least, coordinated by the same agency. This provides a consistent channel of communication for the respondents and represents a cost saving over the two data collection efforts. However, this represents an organisational risk as it is unlikely that one bid will have all the capabilities demanded.

7 REFERENCES

- Abrantes, PAL, Wardman, MR (2011). Meta-analysis of UK values of travel time: An update, *Transportation Research A*, **45**, 1–17.
- Atasoy, D, Glerum, A, Bierlaire, M (2013). Attitudes towards mode choice in Switzerland, *The Planning Review*, 49, 101–117.
- Auger, P, Devinney, TM, Louviere, JJ (2007). Using best–worst scaling methodology to investigate consumer ethical beliefs across countries, *Journal of Business Ethics,* **70**(3), 299-326.
- Bamberg, S, Schmidt, P (2003). Incentives, morality, or habit. Predicting students' car use for university routes with the models of Ajzen, Schwartz, and Triandis, *Environment and Behavior*, **35**, 264–285.
- Ben-Akiva, M, McFadden, D, Train, K, Walker, J, Bhat, C, Bierlaire, M, Bolduc, D, Boersch-Supan, A, Brownstone, D, Bunch, DS, Daly, A, de Palma, A, Gopinath, D, Kalstrom, A, Munizaga, MA (2002). Hybrid Choice Models: Progress and Challenges, *Marketing Letters* **13**(3), 163-175.
- Ben-Akiva, M, Walker, J, Bernardino, A, Gopinath, D, Morikawa, T, Polydoropoulou, A (2002). Integration of Choice and Latent Variable Models. in H Mahmassani (Ed.) *In Perpetual Motion: Travel Behavior Research Opportunities and Application Challenges*, Elsevier Science, 431-470.
- Bliemer, MCJ, Rose, JM, Hensher, DA (2009a). Efficient stated choice experiments for estimating nested logit models, *Transportation Research B*, **43**, 19-35.
- Bliemer, MCJ, Rose, JM, Hesse, S (2009b). Approximation of Bayesian efficiency in experimental designs, *Journal of Choice Modelling*, **1**(1), 98-127.
- Bolduc, D and Alvarez-Daziano, R (2010). On estimation of Hybrid Choice Models in Hess, S, Daly, A (Eds.) *Choice Modelling: The State-of-the-Art and the State-of-Practice*, Emerald, UK, Chp.11, 269-287.
- Bonnel, P, Lee-Gosselin, M, Madre, J-L, Zmud, J (2009). Keeping Up with a Changing World: Challenges in the Design of Transport Survey Methods, in: P Bonnel, MEH Lee-Gosselin, J Zmud and J-L Madre (Eds.) *Transport Survey Methods: Keeping Up with a Changing World*, 3–14. (Bingley: Emerald).
- Bonsall, P (2009). What is so Special about Surveys Designed to Investigate the Environmental sustainability of Travel Behaviour?, in: P Bonnel, MEH Lee-Gosselin, J Zmud and J-L Madre (Eds.) *Transport Survey Methods: Keeping Up with a Changing World*, 49–69. (Bingley: Emerald).
- Brownstone, D, Small, KA (2005). Valuing Time And Reliability: Assessing the Evidence from Road Pricing Demonstrations, *Transportation Research A*, **39**, 279-293.
- Cirillo, C, Axhausen, KW (2006). Evidence on the Distribution of Values of Travel Time Savings from a Six-Week Diary, *Transportation Research A*, **40**(5), 444-457.
- Cohan, H, Southworth, F (1999). On the Measurement and Valuation of Travel Time Variability Due to Incidents on Freeways, *Journal of Transportation and Statistics*, Bureau of Transportation Statistics, 123-131, www.gcu.pdx.edu/download/2cohen.pdf.
- Dagsvik, JK, Wennemo, T, Wetterwald, DG, Aaberge, R (2002). Potential demand for alternative fuel vehicles, *Transportation Research B*, **36**, 361-384.
- De Jong, G, Daly, AJ, Pieters, M, Vellay, C, Hofman, F (2003). A model for time of day and mode choice using error components logit, *Transportation Research E* **39**(3), 245-268.
- Debrezion, G, Pels, E, Rietveld, P (2009). Modelling the joint access mode and railway station choice, *Transportation Research E*, **45**(1), 270-283.

- Department of Transport, Western Australia (2010). *TravelSmart 2010: A 10 year plan*, available at http://www.transport.wa.gov.au/mediaFiles/active-transport/AT_TS_P_tenyearplan.pdf.
- Department of Transport, Western Australia (2014). *TravelSmart*, Perth, Australia, available at http://www.transport.wa.gov.au/activetransport/24605.asp.
- Douglas, NJ, Franzmann, LJ, Frost, TW (2003). Estimation of Demand Parameters for Primary Public Transport Service, Paper presented at the 26th *Australian Transport Research Forum*, October 1-3, Wellington, NZ, available at http://www.atrf.info/papers/2003/2003_douglas_franzmann_frost.pdf.
- Duncan, M (2010). To park or to develop: trade-off in rail transit passenger demand, *Journal of Planning Education and Research* **30**, 162-181.
- Espino, R, Ortúzar, JdD, (2007). Understanding suburban travel demand: Flexible modelling with revealed and stated choice data, *Transportation Research A*, **41**(10), 899-912.
- Ettema, D, Arentze, TA, Timmermans, HJP (2011). Social influences on household location, mobility and activity choice in integrated micro-simulation models, *Transportation Research A*, **45,** 283–295.
- Ewing, G, Sarigollu, E (2000). Assessing consumer preferences for clean-fuel vehicles: a discrete choice experiment, *Journal of Public Policy and Marketing* **19**, 106–118.
- Fernandez-Heredia, A, Monzon, A, Jara-Diaz, S (2014). Understanding cyclists' perceptions, keys for a successful bicycle promotion, *Transportation Research A*, **63**, 1-11.
- Flynn, TN, Louviere, JJ, Peters, TJ, Coast, J (2007). Best–worst scaling: what it can do for health care research and how to do it, *Journal of Health Economics,* **26**(1), 171-189.
- Gatersleben, B, Appleton, KM (2007). Contemplating cycling to work: Attitudes and perceptions in different stages of change, *Transportation Research A*, **41**, 302–312.
- Goodman, R (2001). A Traveller in Time: Understanding Deterrents to Walking to Work, *World Transport Policy and Practice*, **7**(4), 50-54.
- Heffner, R, Kurani, K, Turrentine, T (2007). Symbolism in California's early market for hybrid electric vehicles, *Transportation Research D*, **12**, 396–413.
- Hensher, DA, King, J (2001). Parking demand and responsiveness to supply, pricing and location in the Sydney central business district, *Transportation Research A* **35**(3), 177-196.
- Hensher, DA, Rose, JM, Greene, WH (2005). *Applied Choice Analysis: A Primer*, Cambridge University Press, UK.
- Hess, S, Daly, A, Rohr, C, Hyman, G (2007). On the development of time period and mode choice models for use in large scale modelling forecasting systems, *Transportation Research A* **41**(9), 802-826.
- Hollander, Y (2006). Direct Versus Indirect Models for the Effects of Unreliability, *Transportation Research A*, **40**(9), 699-711.
- Horn, MET (2002). Multi-modal and demand-responsive passenger transport systems: a modelling framework with embedded control systems, *Transport Management*, **36**, 167-188.
- Jensen, AF, Cherchi, E, Mabit, SL (2013). On the stability of preferences and attitudes before and after experiencing an electric vehicle, *Transportation Research D* **25**, 24-32.
- Jensen, M (1999). Passion and heart in transport – a sociological analysis on transport behaviour, *Transport Policy*, **6**, 19–33.
- Kessels, R, Jones, B, Goos, P and Vandebroek, M (2008). Recommendations on the use of Bayesian optimal designs for choice experiments, *Quality and Reliability Engineering International*, **24**(6) 717-744, Special Issue: The European Network for Business and Industrial Statistics (ENBIS).
- Kim, KK, Shin, HK, Kim, B (2011). The role of psychological traits and social factors in using new mobile communication services, *Electronic Commerce Research and Applications* **10**, 408–417.

- Kim, S, Ulfarsson, G, Hennessy, J (2007). Analysis of light rail rider travel behavior: impacts of individual, built environment, and crime characteristics on transit access, *Transportation Research A* **41**, 511-522.
- Krizek, KJ, Roland, RW (2005). What is at the end of the road? Understanding discontinuities of on-street bicycle lanes in urban settings, *Transportation Research D*, **10**, 55–68.
- Lam, TC, Small, KA (2001). The value of time and reliability: measurement from a value pricing experiment, *Transportation Research E*, **37**(2), 231-251.
- Li, Z, Hensher, DA, Rose, JM (2010). Willingness to pay for travel time reliability in passenger transport: A review and some new empirical evidence, *Transportation Research E*, **46**(3), 384-403.
- Li, Z, Tirachini, A, Hensher, DA (2012). Embedding risk attitudes in a scheduling model: application to the study of commuting departure time, *Transportation Science* **46**(2), 170-188.
- Litman, T (2011). Transportation Cost and Benefit Analysis II – Travel Time Costs, Victoria Transport Policy Institute, available at www.vtpi.org/tca/tca0502.pdf.
- Louviere, JJ, Carson, RT, Burgess, L, Street, D, Marley, AAJ (2013). Sequential preference questions factors influencing completion rates and response times using an online panel, *Journal of Choice Modelling*, **8**, 19-31.
- Louviere, JJ, Hensher, DA and Swait, JD (2000). *Stated Choice Methods. Analysis and Application*, Cambridge University Press, UK.
- Mackie, PJ, Wardman, M, Fowkes, AS, Whelan, G, Nellthorp, J, Bates, J (2003). *Values of Travel Time Savings in the UK*, Institute for Transport Studies, University of Leeds, available at http://eprints.whiterose.ac.uk/2079/.
- Marley, AAJ and Islam, T (2012). Conceptual Relations Between Expanded Rank Data and Models of the Unexpanded Rank Data, *Journal of Choice Modelling*, **5(1)**, 38-80.
- McFadden, D (1974). Conditional logit analysis of qualitative choice behaviour in Zarambka, P (Ed.) *Frontiers in Econometrics*, Academic Press, New York, 105-142.
- Metz, D (2008). The Myth of Travel Time Saving, *Transport Reviews*, **28**(3), 321-336, http://pdfserve.informaworld.com/149983 910667966.pdf.
- Meuter, ML, Ostrom, AL, Bitner, MJ, Roundtree, R (2003). The influence of technology anxiety on consumer use and experiences with self-service technologies, *Journal of Business Research*, **56**(11), 899-90.
- Mokhtarian, P, Salomon, I (2001). How Derived is the Demand for Travel? Some Conceptual and Measurement Consideration, *Transportation Research A*, **35**(8), 695-719.
- Murakami, E, O'Connor, S and Gould, J (2009). Electronic Instrument Design and User Interfaces: Synthesis of a Workshop in Bonnel, P, Lee-Gosselin, M, Madre, J-L, and Zmud, J (Eds.) *Transport Survey Methods: Keeping up with a Changing World*, Emerald, UK, 464-472.
- Nutley, SD (1988). 'Unconventional modes' of transport in the United Kingdom—A review of types and the policy context, *Transportation Research A*, **22**(5), 329-344.
- Olaru, D, Smith, B, Wang, J (2011). Optimal discrete choice experimental designs using genetic algorithms, paper presented at the *International Choice Modelling Conference*, Oulton Hall, July 4-6, UK.
- Ozaki, R (2011). Adopting sustainable innovation: What makes consumers sign up to green electricity?, *Business Strategy Environment*, **20**, 1–17.
- Parkhurst, G (2000). Influence of bus-based park and ride facilities on users' car traffic, *Transport Policy*, **7**, 159-172.
- Polydoropoulou, A, Ben-Akiva, M (2001). Combined revealed and stated preference nested logit access and mode choice model for multiple mass transit technologies, *Transportation Research Record* **1771**(1), 38-45.
- Prato, CG, Bekhor, S, Pronello, C (2012). Latent variables and route choice behavior, *Transportation*, 39, 299–319.

- Richardson, AJ, Ampt, ES, Meyburg, AH (1997). Survey Methods for Transport Planning, Eucalyptus Press, Melbourne, available at http://www.geog.ucsb.edu/~deutsch/geog111_211a/code_books/Survey_Methods_For_Transport_Planning.pdf.
- Rose, JM (2013). Interpreting discrete choice models based on best-worst data: A matter of framing, *ITLS WP-13-22*, Institute of Transport and Logistics Studies, The University of Sydney, ISSN 1832-570X.
- Rose, JM and Bliemer, MCJ (2005a). Constructing efficient choice experiments, *Working Paper ITLS WP-05-07*, Institute of Transport and Logistics Studies, The University of Sydney, ISSN 1832-570X.
- Rose, JM and Bliemer, MCJ (2005b). Sample optimality in the design of stated choice experiments, *Working Paper ITLS WP-05-13*, Institute of Transport and Logistics Studies, The University of Sydney, ISSN 1832-570X.
- Rose, JM and Bliemer, MCJ (2009). Constructing efficient stated choice experimental designs, *Transport Reviews*, **29**, 587-617.
- Rose, JM, Bliemer, MCJ, Hensher, DA and Collins, AT (2008). Designing efficient stated choice experiments in the presence of reference alternatives, *Transportation Research E*, **42**, 395-406.
- Ruiz, T, Bernabe, JC (2014). Measuring factors influencing valuation of nonmotorized improvement measures, *Transportation Research A*, **67**(2), 195-214.
- Ryan, M, Gerard, K Amaya-Amaya, M (Eds.) (2007). *Using Discrete Choice Experiments to Value Health and Health Care*, Springer, Dordrecht, The Netherlands.
- Schuitema, G, Anable, J, Skippon, S, Kinnear, N (2013). The role of instrumental, hedonic and symbolic attributes in the intention to adopt electric vehicles, *Transportation Research A*, **48**(2), 39-49.
- Schwanen, T, Mokhtarian, PL (2005). What affects commute mode choice: neighborhood physical structure or preferences toward neighborhoods? Journal of Transport Geography, **13**(1), 83–99.
- Small, KA, Noland, R, Chu, X, Lewis, D (1999). *Valuation of Travel-Time Savings and Predictability in Congested Conditions for Highway User-Cost Estimation*, NCHRP 431, available at http://ntl.bts.gov/lib/16000/16200/16215/PB2000103014.pdf.
- Small, KA, Winston, C, Yan, J (2005). *Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability: Implications for Road Pricing*, University of Irvine, available at www.socsci.uci.edu/~ksmall/Value%20of%20time%20note.pdf.
- Spears, S, Houston, D, Boarnet, MG (2013). Illuminating the unseen in transit use: A framework for examining the effect of attitudes and perceptions on travel behavior, *Transportation Research A*, **58**, 40-53.
- Steg, L (2005). Car use: lust and must. Instrumental, symbolic and affective motives for car use, *Transportation Research A*, **39**, 147-162.
- Taplin, JHE, Qiu, M, Salim, VK, Han, R (2005). *Cost-Benefit Analysis and Evolutionary Computing*, Edward Elgar, Cheltenham, UK.
- Taplin, JHE, Taylor, MAP, Olaru, D, Biermann, S (2014). Transport Modelling Review, Planning and Transport Research Centre (PATREC) Report for the WA Departments of Planning (DoP), Transport (DoT), and MRWA, available at http://www.patrec.org/publication_docs/TMR_Final_report_June2014.pdf.
- Tsamboulas, D, Golias, J, Vlahoyannis, M (1992). Model development for metro station access mode choice, *Transportation* **19**(3), 231-244.
- Van, HT, Choocharukul, K, Fujii, S (2014). The effect of attitudes toward cars and public transportation on behavioral intention in commuting mode choice—A comparison across six Asian countries, *Transportation Research A*,

- Vandenbulcke, G, Dujardin, C, Thomas, I, de Geus, B, Degraeuwe, B, Meeusen, R, Int Panis, L (2011). Cycle commuting in Belgium: Spatial determinants and 're-cycling' strategies, *Transportation Research A*, 45(2), 118-137.
- Venkatesh, V, Davis, FD (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies, *Management Science*, **46**(2), 186-204.
- Vredin Johansson, M, Heldt, T, Johansson, P (2006). The effects of attitudes and personality traits on mode choice, *Transportation Research A*, **40**,507–525.
- Walker, J, Ehlers, E, Banerjee, I, Dugundji, ER (2011) Correcting for endogeneity in behavioral choice models with social influence variables, *Transport Research A*, **45**, 362–374.
- Walker, JL, Li, J (2007). Latent Lifestyle Preferences and Household Location Decisions, *Journal of Geographical Systems*, **9**(1), 77-101.
- Wardman, M (2004). Public Transport Values of Time, *Transport Policy*, **11**(4), 363-377.
- Wen, C-H, Lai, S (2010). Latent class models of international air carrier choice, *Transportation Research E,* **46**(2), 211-221.
- Wen, C-H, Wang, W, Fu, C (2011). Latent class nested logit model for analyzing high-speed rail access mode choice, *Transportation Research E, issue, pp.?*.
- Wener, R, Evans, GW, Lutin, J (2006). *Leave the Driving to Them: Comparing Stress of Car and Train Commuters*, American Public Transportation Association, available at www.apta.com/passenger_transport/thisweek/documents/driving_stress.pdf.
- Yang, K (2012). Consumer technology traits in determining mobile shopping adoption: An application of the extended theory of planned behaviour, *Journal of Retailing and Consumer Services* **19**, 484-491.
- Zamparini, L, Reggiani, A (2007). Meta-Analysis and the Value of Travel Time Savings: A Transatlantic Perspective in Passenger Transport, *Networks and Spatial Economics*, available at www.springerlink.com/content/h663q51u448078x1/fulltext.pdf.
- Ziegler, A (2010). *Individual Characteristics and Stated Preferences for Alternative Energy Sources and Propulsion Technologies in Vehicles: A Discrete Choice Analysis*, Economics Working Paper 10/125, available at http://www.cer.ethz.ch/research/WP-10-125.pdf.

8 APPENDIX A ATTRIBUTES FOR EXPERIMENTAL DESIGNS

Table A1: Attributes

Focus of the design	Attributes	Travel modes	Literature support
Mode Choice	• Travel time • Waiting time • Transfer time • Travel cost or fare • Access and egress times and costs • Frequency services • Travel reliability/congestion • Perceived travel time • Comfort, (in)security • Departure time	All motorised modes	Cohan and Southworth 1999; Mokhtarian and Salomon 2001; Wardman 2004; Wener *et al.* 2006; Metz 2008; Litman 2011 Lam and Small 2001; Small *et al.* 2005; Hollander 2006; De Jong *et al.* 2003; Polydoropoulou and Ben-Akiva 2001
	• Travel time[11] • Infrastructure facilities • Weather	Active transport	Goodman 2001; Litman 2011
Parking charges and regulation	• All attributes for mode choice, plus • Parking costs • Parking penalties • Parking restrictions (i.e. two hour limit) • On street and off street charges • Limited parking space and competition for parking • Time-of-day[12]	All modes may be considered. The trip under investigation is driving to an activity centre (e.g. Perth CBD, major hospital, university).	Parkhurst 2000; Polydoropoulou and Ben-Akiva 2001; Hensher and King 2001; Duncan 2010; Li *et al.* 2010 Debrezion *et al.* 2009 Hess *et al.* 2007; De Jong *et al.*

[11] Under pleasant conditions, walking, cycling, even waiting can have positive values (Goodman 2001).

Focus of the design	Attributes	Travel modes	Literature support
			2003; Li *et al.* 2012
Public transport demand	• All attributes for mode choice, plus • Crowding • Quality of service (headway, frequency) • Having to allow services go by because they are too full • Access modes • Flexible fares encouraging trips outside the peak • Number transfers And for Bus operations: • Fewer bus stops • Integrating buses with rail • Park-and-Ride bus stations (BnR) Or for Rail operations: • Quality of the station facilities • PnR facilities • Patterned services during the peak.	Train, bus, and their combinations versus active transport and car travel	Hensher *et al.* 2011a; Li and Hensher 2013; Tirachini *et al.* 2013; Spears *et al*. 2013 Tsamboulas et al. 1992; Shiftan et al. 2003; Espino *et al*. 2007 Bhatta and Larsen (2011) Parkhurst 2000
Road Demand	Travel time reliability, time spent in congestion, tolled routes, toll charges for infrastructure capacity, congestion charging, managed freeway on ramps	Car focused, but public transport and active transport modes are important	Noland and Small (1995); Hensher *et al.* (2011); Li *et al.* (2012)
Mono-Mode, e.g. Cycling	Distance, separated pathway, perceived safety, facilities at the destination, quality of land-use and natural environment	Cycling focused	Jensen 1999; Krizek and Roland 2005; Gatersleben and Appleton 2007; Vandenbulcke *et al.* 2011; Ruiz and Bernabe 2014; Fernandez-Heredia *et al*. 2014

[12] When considering parking competition, time-of-day becomes an important consideration. The choice set may include travel by same mode, but either earlier or later.

Focus of the design	Attributes	Travel modes	Literature support
New Alternatives, e.g. Light rail	Travel time reliability, Station quality, PnR facilities, in addition to common attributes (all policies)	Replaces other PT on the corridor, access modes and car	Noland and Small (1995); Bates *et al.*, 2001; Kim *et al.* 2007; Hensher *et al.* (2011b); Li *et al.* (2012); Wen *et al.* 2011
Demand responsive transport	Cost of flexible transport, mechanism to express demand, level of flexibility in routing and timing	Flexible mode, bus, taxi, car	Nutley 1988; Horn 2002

9 APPENDIX B EXAMPLES OF ATTITUDINAL QUESTIONS

1) Statements to profile the respondent

Environmental concerns

Ewing and Sarigollu 2000; Dagsvik *et al.* 2002; Heffner *et al.* 2007; Bolduc *et al.* 2008; Hidrue *et al*. 2011; Ozaki 2011; Kim et al. 2011; Schuitema et al. 2013; Fernandez-Heredia *et al*. 2014

"Saving the environment requires our immediate efforts."
"Now is the real time to worry about the effects of air pollution."
"Vehicle emissions can destroy our flora and fauna."
"I am concerned that future generations may not be able to enjoy the world as we know it currently."
"I am willing to spend extra time only to save the environment."
"I always recycle products such as: paper, glass, aluminium, etc."
"I am willing to pay more for products or services only to save the environment."
"I prefer to walk/cycle in order to reduce pollution and greenhouse gases."
"Cycling hardly makes any noise."
"I might join a group, club, or organisation concerned with ecological issues."
"It is acceptable for a modern society to produce a certain degree of pollution."
"Riding public transport helps reduce pollution."
"I prefer driving a car with a powerful engine than a car that emits little CO_2."

Technology adoption

Ewing and Sarigollu, 2000; Parasuraman 2000; Venkatesh and Davis 2000; Meuter *et al.* 2003; Ratchford and Barnhart 2012; Nasri and Charfeddine 2012; Yang 2012

"Using new technologies makes our lives easier."
"Taking up new technologies makes one trendy."
"Things have become so complicated today that it is hard to understand what is going on in this techno-world."
"New technologies cause more problems than they solve."
"I am excited to learn to use new technologies."
"I love gadgets."
"Keeping up with the new knowledge on technologies is necessary."
"I never travel without a GPS."
"People often become too dependent on technology to do things for them."
"I enjoy the challenge of figuring out high-tech gadgets."

Social norms

Venkatesh and Davis 2000; Steg 2005; Ozaki 2011; Yang 2012; Spears *et al.* 2013

"People who influence my behaviour think I should ride a bike more often."
"People who are important to me think that I should ride the bike to work."
"I would go by train if many of my friends would do the same."
"My friends would think it is peculiar not to commute by car."
"My family thinks the problems of car use during rush hours are exaggerated".
"I will not easily travel by bike or bus when all my colleagues drive by car."

Affective factors

Steg 2005; Van *et al.* 2014; Fernandez-Heredia *et al*. 2014
"I express myself through my car."

“The car gives me prestige.”
“I get a kick of driving.”
“I like driving fast.”
“I’m safe in my car.”
“I like my car’s carrying capacity (luggage, purchases).” (4.4.2)
“Riding public transport means you are dependent on others.”
“I can choose my own route.” (4.4.2)
“I take pleasure in riding the bike.”

2) Statements/adjectives to characterise travel modes
Vredin Johansson *et al.* 2006; Jensen *et al.* 2013; Spears *et al.* 2013; Van *et al.* 2014

Richness, luxury, superiority, fashion, and ‘‘coolness’’ - prestige values (each travel mode is described on a scale from poor to rich, or austere to luxurious)
Comfort, excitement, and relaxation - affective feelings
Usefulness, convenience, simplicity, reliability, and speed of the transport mode – instrumental
Environmentally friendly, safe, altruistic, and quiet travel mode – social orderliness

“If I did not need a car, I would dispose of it immediately.”
“The car is always available.”
“I cannot easily drop-off/pick-up passengers without the car.” (4.4.2)
“There are no fuel expenses with cycling.”

10 APPENDIX D EXPERIMENTAL DESIGN TECHNICAL APPENDIX

10.1 BASE DESIGN

Each respondent will see 12 scenarios[13] like the one shown in Figure A1.

Table A3 presents the attribute levels used for trips made at distances less than 20km during morning peak time. The attributes are drawn from literature review. The prior parameters are taken from the Park-and-Ride models undertaken by Huang Ying. The suggested values of waiting time, travel, and access/egress times range between $10 and $15/hour.

Table A3: Attribute levels for the Mode Choice SP experiment

Mode	Attribute	Prior parameter	Levels		
Car only	Expected in-vehicle travel time, TT (min)	-2.1	20	30	40
	Travel time variability, Var (stdev, in min)	-2.4	5	10	15
	Parking cost at the destination, Pcost ($)	-0.3	0	7.5	15
	Fuel/running cost, Fcost ($/100km*d)	-0.2	2.5	3	3.5
	Time-of-day, Driver only vs driver + passengers, (HOV), or any binary attribute relevant to the alternative	0.1		0	1
	ASC car	0.1			
Public transport only	Expected in-vehicle travel time, TT (min)	-1.8	25	30	35
	Travel time variability, Var (stdev, in min)	-2.4	0	2.5	5
	Fare ($)	-0.2	4	4.5	5
	Number of transfers, Trf	-0.2	0	1	2
	Waiting time, WT (min)	-3.0	5	10	15
	Access time, AT (min)	-2.4	4	7	10
	Egress time, ET (min)	-2.4	5	10	15
	ASC public transport	0.05			
Active transport	Expected travel time (min)	-3.6	40	50	60
	Separated footpath/bike lane	0.1		0	1
	Facilities at destination (bike lockers, showers, etc.)	0.2	0	1	2
Combination private-public (e.g., PnR, BnR)	Expected in-vehicle travel time, TT (min)	-1.8	20	30	40
	Travel time variability, Var (stdev, in min)	-2.4	3	5	7
	Parking cost at the station, Pcost ($)	-0.3	2	4	6
	Fuel cost, Fcost ($)	-0.2	0.5	1	1.5
	Time-of-day, Driver only vs driver + passengers, (HOV), or any binary attribute relevant to the alternative	0.1		0	1
	Fare ($)	-0.2	4	4.5	5
	Waiting time, WT (min)	-3.0	5	10	15
	Access time, AT (min)	-2.4	3	5	7
	Egress time, ET (min)	-2.4	5	10	15
	ASC combination private + public transport	0.05			

[13] If they see only six scenarios, the sample sizes below need to be doubled.

A full factorial design with the attributes and levels presented in Table A3 would require more than 80 billion scenarios (83,682,825,624).

Tables A4 and A5 require only a fraction of the full factorial design. While Table A4 provides an optimal design for min Dp error (jointly considering the elements of the AVC matrix), Table A5 focuses only on the variances of the coefficients of interest and tries to minimise the sample size necessary to reliably calculate the most difficult parameter to estimate. For benchmarking purposes and only as an illustration, we also determined the Dz error, to assess the sample size if no information at all is available for the prior parameters (hence considered to be equal to zero). We note that the Fisher Information and the Asymptotic Variance-Covariance Matrices do not include the alternative specific constants (ASC), although they are considered in the calculation of probabilities.

Table A4: Experimental design for the smallest Dp error (Evolver 30 min solution)

Car only					PT							AT			Combination private + public								
x11	x12	x13	x14	x15	x21	x22	x23	x24	x25	x26	x27	x31	x32	x33	x41	x42	x43	x44	x45	x46	x47	x48	x49
30	15	7.5	2.5	1	35	5	4.5	0	15	4	15	50	0	1	20	7	6	0.5	0	4.5	15	3	15
40	10	0	3.5	1	35	0	5	2	5	4	5	40	0	0	30	7	2	1	0	5	5	7	5
20	15	0	3.5	0	35	0	4	1	10	10	5	60	0	0	20	5	2	0.5	1	4	10	3	5
20	15	7.5	3	0	25	0	5	0	5	4	15	60	0	0	40	7	6	0.5	1	5	5	7	15
20	15	0	3	0	25	0	5	0	15	7	15	50	1	2	40	3	2	1.5	1	5	15	7	15
30	15	0	3.5	0	25	5	4.5	0	10	10	5	60	1	1	30	5	4	1	1	4.5	10	5	5
40	5	0	3.5	1	30	5	5	1	10	10	10	60	1	0	40	3	2	0.5	0	5	10	5	10
20	10	0	2.5	1	25	2.5	4	0	15	4	5	60	0	0	20	3	4	1.5	0	4	15	7	5
40	15	0	3.5	1	25	2.5	5	2	5	4	10	40	0	2	40	3	2	0.5	0	5	5	3	10
20	5	15	2.5	0	35	5	4	0	10	4	5	60	1	1	30	7	4	1.5	0	4	10	3	5
40	5	7.5	2.5	0	30	2.5	4	2	15	7	15	50	1	0	40	3	2	1.5	0	4	15	5	15
40	5	15	2.5	0	25	0	4	1	5	10	15	50	1	0	20	3	2	0.5	1	4	5	3	15

Table A5: Experimental design for smallest sample size (Evolver 30 min solution)

Car only					PT							AT			Combination private + public								
x11	x12	x13	x14	x15	x21	x22	x23	x24	x25	x26	x27	x31	x32	x33	x41	x42	x43	x44	x45	x46	x47	x48	x49
30	15	7.5	2.5	1	35	5	4.5	0	10	4	15	50	0	1	20	7	6	0.5	0	4.5	10	3	15
30	10	0	3.5	1	35	0	5	2	5	4	5	40	0	0	30	5	2	1	0	5	5	7	5
20	15	0	3.5	0	35	2.5	4	1	10	10	15	60	0	0	20	5	2	0.5	1	4	10	3	15
20	10	7.5	3	0	25	5	5	0	5	4	10	60	0	0	40	7	6	0.5	1	5	5	7	10
20	15	0	3	0	30	0	5	0	15	7	15	60	1	2	40	3	2	1.5	1	5	15	7	15
30	15	0	3.5	0	30	5	4.5	0	10	10	5	60	1	1	30	5	4	1	1	4.5	10	5	5
40	5	0	3.5	1	30	5	5	1	15	10	10	60	1	0	30	5	2	0.5	0	5	15	3	10
30	10	0	2.5	1	30	2.5	4	0	15	7	5	60	0	0	20	3	4	1.5	0	4	15	7	5
30	15	0	3.5	1	25	2.5	5	2	5	4	10	40	0	2	40	3	2	0.5	0	5	5	3	10
20	5	15	2.5	0	35	2.5	4	0	10	4	5	60	1	1	20	7	4	1.5	0	4	10	5	5
40	5	7.5	2.5	0	30	2.5	4	2	5	7	10	50	1	0	30	3	2	1.5	0	4	5	5	10
40	10	15	2.5	0	25	0	4	1	15	10	10	50	1	0	30	3	2	0.5	1	4	15	3	10

Similar to prior work, D_p and S designs outperform orthogonal designs and the S design offers sample sizes an order of magnitude even smaller than the D_p design.
Bliemer and Rose (2005) analysed the impact of the number of choice situations, number of attribute levels, and the attribute range and found best results for 12 experiments, two level attributes with wide level range, both in terms of D-error and sample size S. However, having two levels with wide ranges indices dominance (Bliemer and Rose, 2005: p. 16), which suggests that three levels are preferred. These recommendations were used in the template designs presented in this report.

Bliemer and Rose (2005) have also assessed the role of "wrong" priors and concluded that the designs are quite stable/robust as long as the ratios between the parameters do not change excessively. A test with three sets of prior parameters was undertaken (one presented in Table A3, another one proportional to the values in Table A3 (Table A4), and another one without proportionality – see Table A6), and the results show that when the priors vary considerably, the features of the design also vary. Notably, the "worst" Dp and S values occur for the designs with the "poorer" prior parameter estimates, when their values are arbitrarily chosen, paying attention only to their sign. These results further support the implementation of Bayesian designs, likely to lead to the most robust experimental designs to errors on prior parameters.

Table A4: Summary mode choice SP experiment features and sample sizes for priors proportional to the values in Table A3 (5 times lower, except ASC)

Orthogonal design	Dp-optimised design	S-optimised design
D_p = 28.13 S = 10^9	D_p = 6.231 S_{min} = 20,430	D_p = 10.859 S_{min} = 2,182

Table A5: Prior parameters for the Mode Choice SP experiment non-proportional to the values in Table A3

Mode	Attribute	Prior parameter
Car only	Expected in-vehicle travel time (min)	-0.25
	Travel time variability (stdev, in min)	-0.4
	Parking cost at the destination ($)	-0.06
	Fuel/running cost ($/100km*d)	-0.04
	Time-of-day, Driver only vs driver + passengers, (HOV), or any binary attribute relevant to the alternative	0.1
	ASC car	0.1
Public transport only	Expected in-vehicle travel time (min)	-0.25
	Travel time variability (stdev, in min)	-0.4
	Fare ($)	-0.04
	Number of transfers	-0.04
	Waiting time (min)	-0.2
	Access time (min)	-0.3
	Egress time (min)	-0.3
	ASC public transport	0.05
Active transport	Expected travel time (min)	-0.4
	Separated footpath/bike lane	0.1
	Facilities at destination (bike lockers, showers, etc.)	0.2
Combination private-public (e.g., PnR, BnR)	Expected in-vehicle travel time (min)	-0.1
	Travel time variability (stdev, in min)	-0.4
	Parking cost at station ($)	-0.12
	Fuel cost ($)	-0.04
	Time-of-day, Driver only vs driver + passengers (HOV), or any binary attribute relevant to the alternative	0.1
	Fare ($)	-0.04
	Waiting time (min)	-0.2
	Access time (min)	-0.3
	Egress time (min)	-0.3
	ASC combination private + public transport	0.05

Note: The values are randomly chosen and do not reflect valuations of travel time.

Table A6: Summary mode choice SP experiment features and sample sizes for priors non-proportional to the values in Table 3

Orthogonal design	Dp-optimised design	S-optimised design
D_p = 41.55 S = 10^9	D_p = 9.307 S_{min} = 25,293	D_p = 14.825 S_{min} = 2,986

10.2 LIGHT RAIL

10.2.1 Sample Size Estimate

Table A7: Attribute levels for the Light Rail SP experiment

Mode	Attribute	Levels		
Car only	Expected in-vehicle travel time (min)	20	30	40
	Travel time variability (stdev, in min)	5	10	15
	Parking cost at the destination ($)	0	7.5	15
	Fuel/running cost ($/100km*d)	2.5	3	3.5
	Time-of-day, Driver only vs driver + passengers (HOV), or any binary attribute relevant to the alternative		0	1
	ASC car			
Public transport (Train or Bus as appropriate)	Expected in-vehicle travel time (min)	20	25	30
	Travel time variability (stdev, in min)	0	2.5	5
	Fare ($)	4	4.5	5
	Transfers		0	1
	Waiting time (min)	8	10	12
	Access time (min)	8	10	12
	Egress time (min)	5	10	15
	ASC public transport			
Light rail	Expected in-vehicle travel time (min)	30	35	40
	Travel time variability (stdev, in min)	0	2.5	5
	Fare ($)	5	6	7
	Transfers		0	1
	Waiting time (min)	2	4	6
	Access time (min)	4	6	8
	Egress time (min)	4	8	12
Combination private-public (e.g., PnR, BnR)	Expected in-vehicle travel time (min)	20	30	40
	Travel time variability (stdev, in min)	3	5	7
	Parking cost at station ($)	2	4	6
	Fuel cost ($)	0.5	1	1.5
	Time-of-day, Driver only vs driver + passengers, (HOV), or any binary attribute relevant to the alternative		0	1
	Fare ($)	4	4.5	5
	Waiting time (min)	5	10	15
	Access time (min)	3	5	7
	Egress time (min)	5	10	15
	ASC combination private + public transport			

The total number of the combinations for a full factorial design with the attributes and levels presented in Table 13 would require 1.067*10^13 scenarios.

Table A8: Experimental design for the smallest Dp error (Evolver 30 min solution)

Car only					PT							LR							Combination private + public								
x11	x12	x13	x14	x15	x21	x22	x23	x24	x25	x26	x27	X31	X32	X33	X34	X35	X36	X37	x41	x42	x43	x44	x45	x46	x47	x48	x49
40	15	15	2.5	0	20	5	5	1	8	8	15	35	0	7	0	6	4	8	30	3	4	0.5	1	5	10	3	15
20	5	7.5	2.5	0	30	0	4	0	10	8	5	40	0	7	1	4	8	12	40	7	2	0.5	0	4.5	10	3	5
40	15	0	2.5	0	25	0	4	2	12	8	15	30	5	6	1	2	8	4	20	5	6	1.5	0	5	5	3	10
30	15	7.5	3.5	1	30	0	4.5	2	8	10	10	35	5	5	0	6	6	12	40	7	6	1	1	4	5	5	5
40	10	7.5	2.5	0	25	2.5	4	0	12	12	15	35	2.5	5	1	6	6	8	20	7	6	0.5	1	4	10	7	10
20	5	15	3	1	30	5	4	1	8	10	15	40	0	6	2	2	8	8	40	5	4	1.5	1	4	5	7	10
30	15	7.5	3.5	0	30	5	5	1	12	8	5	40	5	5	0	6	8	4	20	3	4	1.5	0	4	15	7	10
20	15	15	2.5	0	25	5	4	0	8	12	5	30	5	7	2	4	4	4	40	3	2	0.5	1	4.5	15	5	15
40	10	7.5	3.5	1	20	5	4	2	12	12	5	30	5	5	2	2	4	12	30	7	2	1.5	1	4	15	3	5
40	15	0	3	1	20	0	4.5	0	10	8	10	30	0	5	2	6	8	4	20	3	2	1.5	0	4.5	15	5	15
20	15	0	3.5	0	20	5	5	1	8	12	5	40	5	6	0	2	4	4	20	7	2	1	1	5	5	7	15
40	5	0	3.5	1	30	2.5	4	2	12	12	5	40	2.5	5	2	6	4	4	20	3	2	0.5	0	5	5	5	10

Table A9: Experimental design for smallest sample size (Evolver 30 min solution)

Car only					PT							LR							Combination private + public								
x11	x12	x13	x14	x15	x21	x22	x23	x24	x25	x26	x27	X31	X32	X33	X34	X35	X36	X37	x41	x42	x43	x44	x45	x46	x47	x48	x49
40	15	15	2.5	0	25	2.5	5	2	8	8	15	35	0	7	1	6	4	8	30	3	4	1	1	5	10	3	15
20	5	7.5	2.5	0	30	0	4.5	0	10	8	5	40	0	7	1	4	8	12	40	7	4	0.5	0	4.5	10	3	5
40	15	0	2.5	0	25	0	4	2	12	12	15	30	2.5	6	1	2	8	4	20	5	6	1.5	0	5	5	3	10
30	15	7.5	3.5	1	30	0	4.5	1	8	10	10	35	5	5	0	6	6	12	40	7	6	1	1	4	5	5	5
40	10	7.5	3.5	0	25	2.5	4	0	12	12	15	35	2.5	5	1	6	6	8	20	7	6	0.5	1	4	10	7	15
20	5	15	3	1	30	5	4	1	8	10	15	40	0	6	2	2	8	8	40	5	4	1.5	1	4	5	7	10
30	15	15	3	0	30	5	5	1	12	8	5	40	5	5	0	6	8	8	20	3	4	1.5	0	5	15	7	10
20	5	15	2.5	0	25	5	4	0	8	10	5	30	5	6	1	4	4	4	40	3	2	0.5	1	4.5	15	5	15
20	10	7.5	2.5	1	20	5	5	1	12	12	5	30	2.5	5	2	2	6	12	30	7	2	1.5	1	4	15	3	5
40	15	7.5	3	1	20	0	4	0	10	8	10	30	2.5	5	2	6	8	4	20	3	2	1.5	0	4.5	15	5	15
20	15	0	3.5	0	20	5	5	1	8	12	5	40	5	6	1	6	4	4	20	7	2	1	1	5	10	7	15
40	5	0	3.5	1	20	2.5	4	2	12	12	5	30	2.5	6	2	4	4	8	20	3	4	0.5	0	4.5	5	5	10

Table A10: Summary light rail SP experiment features and sample sizes

Orthogonal design	Dp-optimised design	S-optimised design
D_p = 72.94 S = 10^10	D_p = 15.725 (Dz = 13.077) S_{min} = 743 (Sz = 756)	D_p = 19.813 S_{min} = 541

10.3 CAR TOLL

10.3.1 Sample Size Estimate

Table A11: Attribute levels for the Car Toll SP experiment

Mode	Attribute	Prior parameter	Levels		
Car	Expected in-vehicle travel time (min)	-2.1	20	30	40
	Travel time variability (stdev, in min)	-24	5	10	15
	Parking cost at the destination ($)	-0.3	0	7.5	15
	Fuel/running cost ($/100km*d)	-0.2	2.5	3	3.5
	Time-of-day, Driver only vs driver + passengers (HOV), or any binary attribute relevant to the alternative	0.1		0	1

Mode	Attribute	Prior parameter	Levels		
	ASC car	0.1			
Car toll	Expected in-vehicle travel time (min)	-2.1	15	25	35
	Travel time variability (stdev, in min)	-24	3	6	9
	Parking cost at the destination ($)	-0.3	0	7.5	15
	Fuel/running cost ($/100km*d)	-0.2	2	2.5	3
	Toll/congestion charge	-0.2	2	4	6
	Driver only vs driver + passengers (HOV)	0.1		0	1
Public transport only	Expected in-vehicle travel time, *TT* (min)	-1.8	30	35	40
	Travel time variability, *Var* (stdev, in min)	-24	0	2.5	5
	Fare ($)	-0.2	4	4.5	5
	Number of transfers, *Trf*	-0.05	0	1	2
	Waiting time, *WT* (min)	-3	5	10	15
	Access time, *AT* (min)	-2.4	5	10	15
	Egress time, *ET* (min)	-2.4	5	10	15
	ASC public transport	0.05			
Combination private-public (e.g., PnR, BnR)	Expected in-vehicle travel time, *TT* (min)	-1.8	20	25	30
	Travel time variability, *Var* (stdev, in min)	-24	3	5	7
	Parking cost at the station, *Pcost* ($)	-0.3	2	4	6
	Fuel cost, *Fcost* ($)	-0.2	0.5	1	1.5
	Time-of-day, Driver only vs driver + passengers (*HOV*), or any binary attribute relevant to the alternative	0.1		0	1
	Fare ($)	-0.2	4	4.5	5
	Waiting time, *WT* (min)	-3	5	10	15
	Access time, *AT* (min)	-2.4	3	5	7
	Egress time, *ET* (min)	-2.4	5	10	15
	ASC combination private + public transport	0.05			

The total number of the combinations for a full factorial design with the attributes and levels presented in Table A11 would 28 billion (27,894,275,208) scenarios..

Table A12: Experimental design for the smallest Dp error (Evolver 30 min solution)

Car only					Car Toll						PT							Combination private + public								
x1 1	x1 2	x1 3	x1 4	x1 5	x2 1	x2 2	x2 3	x2 4	x2 5	x2 6	X3 1	X3 2	X3 3	X3 4	X3 5	X3 6	X3 7	x4 1	x4 2	x4 3	x4 4	x4 5	x4 6	x4 7	x4 8	x4 9
30	5	7.5	3.5	1	15	3	7.5	6	2.5	1	35	5	5	1	5	15	10	25	7	6	0.5	0	5	5	7	10
20	15	15	3	0	15	9	15	2	2.5	0	40	5	4	2	15	5	15	20	7	4	1.5	0	4	15	7	15
30	5	15	2.5	1	15	3	15	2	2	1	40	0	5	0	15	5	5	30	3	2	0.5	0	5	15	5	5
20	15	7.5	2.5	0	15	3	7.5	6	2	0	30	0	5	2	10	10	15	30	5	4	1.5	1	5	15	7	15
40	15	0	3	0	15	9	0	6	2	0	40	0	4	0	5	10	15	30	7	2	0.5	1	4	5	3	15
40	10	0	3.5	0	25	3	0	4	3	0	40	2.5	5	2	5	5	5	20	3	2	1	1	5	10	5	5
40	5	7.5	3.5	0	35	3	7.5	6	3	0	40	2.5	4	2	15	15	15	30	3	6	1.5	0	4	10	3	15
20	15	7.5	3.5	0	15	3	7.5	2	2	0	30	5	5	0	15	15	10	20	3	6	1	1	5	5	3	10
40	10	0	3.5	1	35	6	0	4	2	1	30	0	4.5	1	15	5	15	20	3	2	0.5	0	4.5	5	7	15
20	15	7.5	3.5	1	15	9	7.5	2	3	1	35	0	4.5	2	15	15	5	25	5	2	1.5	0	4.5	5	3	5
40	15	7.5	2.5	1	25	9	7.5	6	2	1	40	5	4	2	15	15	5	30	3	6	0.5	1	4	5	7	5
40	15	7.5	2.5	0	35	6	7.5	2	2	0	40	0	5	2	10	15	5	20	7	6	0.5	0	5	15	3	5

Table A13: Experimental design for smallest sample size (Evolver 30 min solution)

Car only					Car Toll						PT							Combination private + public								
x11	x12	x13	x14	x15	x21	x22	x23	x24	x25	x26	X31	X32	X33	X34	X35	X36	X37	x41	x42	x43	x44	x45	x46	x47	x48	x49
30	10	7.5	3.5	1	15	3	7.5	4	2.5	0	35	5	5	1	5	15	10	25	7	4	1	1	5	5	7	10
20	15	15	3	0	15	6	15	2	2.5	0	40	5	4	2	10	15	15	20	7	4	1.5	0	4	5	7	15
30	5	15	2.5	1	15	3	15	2	2	0	40	0	5	0	15	5	10	30	3	4	0.5	1	5	15	5	10
40	15	7.5	3	0	15	3	7.5	6	2.5	1	30	0	5	2	10	10	15	30	5	4	1.5	0	5	15	7	15
40	15	0	3	0	15	9	0	6	2.5	1	40	0	4	0	5	10	15	25	7	2	0.5	0	4	5	3	15
40	10	0	3.5	0	25	6	0	4	3	1	40	2.5	4.5	2	5	5	5	20	3	2	1	0	4.5	10	5	5
40	15	7.5	3.5	0	35	6	7.5	4	3	0	40	2.5	5	2	15	15	10	30	3	6	1	0	5	10	7	10
20	5	7.5	3.5	0	15	3	7.5	2	2	1	35	5	5	0	15	10	10	25	5	4	1.5	0	5	5	3	10
40	10	0	2.5	1	35	6	0	4	2	0	30	0	4.5	1	5	5	15	20	3	2	1.5	1	4.5	5	3	15
20	15	7.5	3.5	1	15	3	7.5	2	3	0	35	0	4.5	2	15	10	5	25	5	4	1.5	1	4.5	10	3	5
40	15	7.5	3.5	1	25	9	7.5	6	2	1	40	5	4	2	15	15	10	20	5	6	1	1	4	10	7	10
40	15	7.5	3.5	0	35	6	7.5	4	2	0	40	5	5	2	10	15	5	20	7	4	1	0	5	15	3	5

Table A14: Summary Car toll SP experiment features and sample sizes

Orthogonal design	Dp-optimised design	S-optimised design
D_p = 31.72 S = 10^9	D_p = 6.794 (Dz =) S_{min} = 1,564 (Sz =)	D_p = 9.423 S_{min} = 959

It is important to note that this design has several additional constraints:

- The public transport fare and the egress time are considered the same for the Public transport only and for Combination PT and private;
- Parking cost at destination is the same for the two private transport alternatives car and car on a tolled route;
- Car travel time and associated fuel cost must be larger than the travel time and fuel cost on the tolled route; similarly for variation of travel time.

These constraints affect the efficiency of the design and consequently, require larger sample sizes.

11 APPENDIX D EXAMPLES SCENARIOS

Figure A1: Example scenario for Mode Choice baseline SP (Table 2)

Car

Travel time

Travel Time	30 min
Travel Time Variation	15 min

Cost of trip

Fuel Cost	$2.50
Parking Cost at Destination	$7.50

Public Transport

Travel time

Access Time	4 min
Waiting Time	10 min
In-vehicle Travel Time	35 min
In-vehicle Travel Time Variation	5 min
No. of Transfers	0
Egress Time	15 min

Cost of trip

Public Transport Fare	$4.50

Car + Public Transport

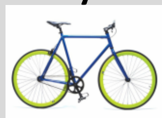
Travel time

Access Time	3 min
Waiting Time	10 min
In-vehicle Travel Time	20 min
In-vehicle Travel Time Variation	7 min
Egress Time	15 min

Cost of trip

Public Transport Fare	$4.50
Parking Cost at Station	$6
Fuel Cost for Car	$0.50

Bicycle



Travel time

Trip Time	50 min
Bicycle Lane	No separated bike lane
Facility at destination	Bike lockers

Cost of trip

Trip Cost	$0

Best

Worst

Figure A2: Example scenario for financial instruments to manage congestion (Table 6)

Car

Travel time

Travel Time	30 min
Travel Time Variation	10 min

Cost of trip

Fuel Cost	$3.50
Parking Cost at Destination	$7.50

Car

Travel time

Travel Time	15 min
Travel Time Variation	3 min

Cost of trip

Fuel Cost	$2.50
Parking Cost at Destination	$7.50
Toll	$4

Public Transport

Travel time

Access Time	15 min
Waiting Time	5 min
In-vehicle Travel Time	35 min
In-vehicle Travel Time Variation	5 min
No. of Transfers	1
Egress Time	10 min

Cost of trip

Public Transport Fare	$5

Car + Public Transport

Travel time

Access Time	7 min
Waiting Time	5 min
In-vehicle Travel Time	25 min
In-vehicle Travel Time Variation	7 min
Egress Time	10 min

Cost of trip

Public Transport Fare	$5
Parking Cost at Station	$4
Fuel Cost for Car	$1

Please indicate below your most preferred and the least preferred travel options out of the four alternatives:

	Car	Car	Public Transport	Car + Public Transport
Best				
Worst				

Figure A3: Example scenario Light Rail design (Table 4)

Car

Travel time

Travel Time	40 min
Travel Time Variation	15 min

Cost of trip

Fuel Cost	$2.50
Parking Cost at Destination	$15

Public Transport

Travel time

Access Time	8 min
Waiting Time	8 min
In-vehicle Travel Time	25 min
In-vehicle Travel Time Variation	2.50 min
No. of Transfers	2
Egress Time	15 min

Cost of trip

Public Transport Fare	$5

Light Rail

Travel time

Access Time	4 min
Waiting Time	6 min
In-vehicle Travel Time	35 min
In-vehicle Travel Time Variation	0 min
No. of Transfers	1
Egress Time	8 min

Cost of trip

Public Transport Fare	$7

Car + Public Transport

Travel time

Access Time	3 min
Waiting Time	10 min
In-vehicle Travel Time	30 min
In-vehicle Travel Time Variation	3 min
Egress Time	15 min

Cost of trip

Public Transport Fare	$5
Parking Cost at Station	$4
Fuel Cost for Car	$1

Please indicate below your most preferred and the least preferred travel options out of the four alternatives:

Best

Worst